What Happens Once the Intervention Ends?
The Five-Year Impacts of a Cash Transfer Experiment in Malawi

Sarah Baird[*], George Washington University
Ephraim Chirwa, University of Malawi, Chancellor College
Craig McIntosh, University of California, San Diego
Berk Ozler, The World Bank
October 26, 2014

FINAL REPORT

Abstract

This report evaluates the five year impacts of the Zomba Cash Transfer Program (ZCTP) in Zomba, Malawi. The ZCTP took place for two years during 2008-2009, and involved giving cash transfers, both conditional on schooling and unconditionally, to initially never-married 13-22 year old young women. The Schooling, Income and Health Risk (SIHR) study was designed to evaluate the impacts of the cash transfer program on a variety of outcomes ranging from education to health to sexual behavior. The SIHR study is a randomized control trial where young women were randomly assigned to one of three groups: control, Unconditional Cash Transfer (UCT), and Conditional Cash Transfer (CCT). Baseline data was collected in 2007, with follow-up data collected in 2008-2009 (Round 2), during the program, in 2010 (Round 3), immediately upon the conclusion of the program and in 2012-2013 (Round 4), two years after the program ended.

The strong and significant short-term impacts of the ZCTP (using data collected in 2008 and 2010) have been documented elsewhere. This report focuses on impacts two years after the program ended, in 2012-2013, to try and understand whether this relatively short (two-year) intervention of cash transfers – introduced at a particularly important period of transition from adolescence to adulthood – can have lasting effects on this cohort of young females and their future families. The analysis focuses on four key domains for the recipients of the cash transfer program: education; marriage and fertility; health and nutrition; and sexual behavior. The analysis focuses on whether results found in the short term were sustained two years after the program ended. Results are analyzed separately for young women who were in school at baseline (*baseline schoolgirls*) and those that were out of school at baseline (*baseline dropouts*), an oft-overlooked group. The analysis for baseline schoolgirls focuses on differential impacts between the CCT and UCT arm, while the analysis for baseline dropouts focuses on the difference between the CCT and the control (no UCT experiment was conducted for this group). Overall, results suggest that the substantial benefits conferred by unconditional treatment while the program was in place in the domains that we investigate here were almost completely transient. Even the conditional program, when implemented among those in school at baseline (and therefore likely to continue with schooling even in the absence of a CCT) had few detectable long-term impacts. The program that provided

---

[*] Corresponding author: sbaird@gwu.edu. Research discussed in this publication has been funded by the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN). The views expressed in this article are not necessarily those of 3ie or its members. Funding was also received by the Research Support Budget of the World Bank.

conditional cash transfers to girls who had already dropped out of school at baseline, on the other hand, had large and durable impacts on a wide range of outcomes – including primary school completion, years of education, marriage rates, likelihood of having started childbirth, and desired fertility. Our results suggest that long-term impacts are sustained only when a cash transfer program achieves substantial improvements in the stock of a durable form of capital, such as human capital. The results of the evaluation of the ZCTP, both in the short and long term, provide important lessons for policy makers thinking of designing cash transfer programs as part of their social protection policy.

Table of Contents

1. Introduction

Adolescent girls in sub-Saharan Africa (SSA) face a multitude of hazards during their transition from childhood to adulthood. Net primary school enrollment for females in the region is lower than 75% -- with lower completion rates and much lower transition rates into secondary school, resulting in net enrollment rates in secondary schools of around 25% in countries like Malawi (World Bank 2013). Age at first marriage, while recently increasing, remains around 18 or 19 in many countries in the region (Marston et al. 2009; Garenne 2008) – with adolescent childbearing rates showing no decline since the 1990s and remaining higher than 100 births per 1,000 adolescent girls in nearly half of the countries (UNFPA 2012). Onset of sexual activity at an early age with older male partners exposes young women to sexually transmitted diseases, with HIV prevalence among females aged 15-24 at 22.7% in Swaziland in 2007, 13.6 in South Africa in 2009 and 5.2 in Malawi in 2010. Along with HIV/AIDS and abortion, depression makes up the leading contribution to disability adjusted life-years in SSA, with one recent study in Malawi showing that more than a third of school-aged girls suffering from psychological distress (Patel et al. 2007; Baird, de Hoop, and Özler 2013).

Governments and aid organizations have responded to this dire picture by designing a variety of interventions targeting school-aged girls and young women. For example, donor organizations like DFID have formed strategic collaborations with foundations like Nike and NoVo to create the Girl Hub and invest in policies and programs that benefit adolescent girls and will have lasting impacts (http://www.girleffect.org/about/girl-hub/). Such interventions are wide ranging in their approaches. A recent review of interventions targeted at adolescents in low- and middle-income countries that reported effects on childbearing-related outcomes lists five types of interventions: cash transfer programs, communications, peer education, school-based interventions or workshops, and health services or counseling (McQueston, Silverman, and Glassman 2013).[1]

Rigorous evaluations of these interventions generally indicate positive, albeit modest, effects. A recent systematic review of cash transfer programs by Baird et al. (2013) indicates that both conditional and unconditional cash transfer programs improve school enrollment and attendance, with little effect on test scores – generally confirming findings from earlier reviews of conditional cash transfer programs (Fiszbein and Schady 2009; Saavedra and Garcia 2012). Short-term evidence from the cash transfer experiment under examination in this paper shows that cash transfers can also significantly delay marriage and pregnancy among school-aged girls (Baird et al. 2010; Baird, McIntosh, and Özler 2011); reduce the prevalence of HIV and HSV-2 (Baird et al. 2012); and improve psychological wellbeing (Baird, de Hoop, and Özler 2013). The evaluation of a school-based intervention in Kenya testing the effects of education subsidies and HIV education separately and jointly found significant reductions in school dropout, pregnancy, and marriage among girls in the stand-alone education subsidy arm, and a modest reduction in HSV-2 in the joint program arm (Duflo, Dupas, and Kremer 2012). Empowerment and Livelihood for Adolescents (ELA) in Uganda showed modest declines in childbearing after two years, large reductions in having had sex unwillingly, and increases in self-employment activities.

[1] Other interventions targeted at this group are variants of interventions in these broad categories, such as the Empowerment and Livelihood for Adolescents (ELA) in Uganda that provides life skills and vocational skills through mentors in adolescent development clubs (Bandiera et al. 2012); peer-led sessions in Safe Spaces in Bangladesh combined with (in-kind) incentives to delay marriage until the legal age of 18 (Field and Glennerster, 2007); etc.

Promising as findings from these studies are, many of them report short-term outcomes that are measured during or immediately after the program: a typical follow-up period would be conducted at 12 or 24 months after baseline.[2] If the aim of these programs is to not only increase current welfare for adolescents, but actually improving their lives in the long-run by making investments in their human and physical capital during an important period of transition in their lives, then it is important to find out whether the short-term improvements were temporary or sustained. The welfare of these adolescents as adults – as well as their families – will improve only if the interventions altered their life trajectories.

This study aims to help fill this gap by reporting outcomes over a five-year timespan for a cash transfer experiment that ran for two years, from 2008-2010. The impacts of this program estimated at the midpoint and the end of the experiment have been previously reported. The "Schooling, Income, and Health Risk" (SIHR) project is a randomized prospective study of the Zomba Cash Transfer Program (ZCTP), which was designed to test the importance of key parameters in the design of cash transfer programs. Specifically, SIHR assessed the effects of offering cash transfers to the families of school-age girls for a duration of two years – while randomly assigning key policy design parameters, such as conditionality of transfers on school attendance, transfer amount, and the recipient of the transfers within households (parents or adolescent girl). As described above, the one- and two-year impacts suggest that cash transfers had significant effects on outcomes ranging from education to early marriage and pregnancy to mental health. In this paper, we present the trajectory for the same set of outcomes – from baseline in 2007 until the five-year follow-up in 2012 – separately for girls who received cash transfers conditional on school attendance (CCT), who received transfers unconditionally (UCT), and those who did not receive them (control). This analysis therefore provides a view of how the lives of study participants have evolved two years after the cessation of the program.

Findings from earlier rounds of data collection suggest that there may be good reasons to think that the trajectories of impacts may diverge across the strata and treatment arms of the study. For example, Baird, McIntosh, and Özler (2011) shows that the significantly larger reductions in early marriage and teenage pregnancy in the UCT group is due to an income effect on these outcomes – mostly due to the effect of cash transfers among those who dropped out of school. This indicates that these trends could be reversed once the cash transfers are discontinued. Furthermore, the same study also found larger enrollment and learning effects in the CCT group, who may go on to stay in school longer, achieve higher grade attainment, and delay marriage and childbearing. Hence, an assessment of longer-term outcomes, particularly following the end of the intervention after two years, is likely to give us a more complete picture of temporary vs. sustained effects of CCTs and UCTs.

There is another reason why examining the longer-term outcomes is interesting. While the experimentation with conditionality was conducted among 'baseline schoolgirls,' those in the target population of never-married 13-22 year-old females who reported being enrolled in school at baseline, the smaller group of 'baseline dropouts,' who had dropped out of school at baseline and formed about 15% of the target population, were only offered CCTs. However, the effects of cash transfers on this 'at risk' group were much larger in magnitude than the effects among 'baseline schoolgirls.' For example, at the one-year follow-up, 17% of the control group of baseline dropouts reported having returned to school, compared with 61% of those offered CCTs. Observing a variety of outcomes among this group over a period of five years allows us to examine whether any significant changes followed as a result of the

[2] Duflo, Dupas, and Kremer (2012) is an exception, reporting schooling, marriage, and pregnancy outcomes at three-, five-, and seven-year follow-ups and biological outcomes for sexually transmitted infections at seven-year follow-up.

substantial increases in school enrollment and learning observed during the intervention period.

Our findings may serve to temper some of the recent enthusiasm for unconditional transfers, by showing that the substantial effects we observed while the program was in place were almost entirely transient. Not only did the substantial wedge the UCT program created in critical outcomes like teen pregnancy and early marriage not continue to grow once the program was ended; the UCT arm appears to have suffered from a 'bounce-back' whereby the differential growth rate in these outcomes once the program ended appears to be the mirror image of the short-term treatment effects. This indicates that while the UCT did temporarily delay the onset of negative outcomes for adolescent girls while the program was in place, as soon as it ended the UCT arm engaged in 'catch-up' behaviors that caused the trajectory of these outcomes to return to *the* same level that they would have reached had the program never been put in place, and this within only a very few years. Because the UCT did not lead to the accumulation of any kind of capital, whether human, social, or physical, it appears to have impacts that were very substantial but also entirely transitory.

The SIHR study contained one subgroup that achieved a very meaningful increase in human capital relative to the counterfactual, and this is the group that was out of school as of baseline and then received conditional transfers. In this group the program led to a causal increase of about six tenths of a year of school, and this gap remained almost exactly consistent over the longer term (presumably because the control group was almost entirely out of school, and the treatment cannot 'bounce back' by undoing years of education they acquired during the program). The baseline dropouts' likelihood of completing primary school actually increased between rounds 3 and 4. Because this stock of acquired human capital did not erode relative to the control, we then see a wide range of other behaviors remaining improved over the longer run: they are 10pp less likely to be married, 4 pp less likely to have ever been pregnant, and their number of live births and desired fertility decreased by approximately 0.15. Consistent with other studies we have found no evidence that cash transfers lead to meaningful accumulations of physical capital. Thus one way of framing our results is that the only the subgroup to experience real long-term changes in a variety of outcomes was the one in which cash transfers caused a meaningful accumulation of human capital during and after the two-year program.

The remainder of this report is structured as follows. Section 2 provides a brief review of some of the literature and provides details of the setting of the study. Section 3 discusses the intervention in detail and provides a theory of change and a discussion of the relevant outcome measures. Section 4 provides some additional detail on program implementation. Section 5 presents the results, and Section 6 discusses the policy recommendations.

2. Background, Study Design, and Data[3]

2.1. Relevant Literature

[3] This section draws heavily from McIntosh, Baird, and Özler (2011).

The question of whether behavioral conditions, such as vaccinating children or sending them to school, should be attached to cash transfers remains a highly debated policy topic. While there is little doubt – theoretically and based on empirical evidence – that conditions cause a change in the behavior in question (over and above that would be attained by cash transfers with no strings attached), there are several legitimate objections to attaching conditions to cash transfers – particularly those that form an important part of a government's social safety nets.[4] First, given the higher administrative needs, CCTs are more intensive to run than UCTs. Second, many UCT programs have been found to lead to increases (albeit more muted) in precisely the outcomes on which CCTs are typically conditioned. Finally, an earlier paper in this study was influential in arguing that CCTs may undermine the social protection dimension of cash transfer programs because they create incentives for behavior change precisely by denying transfers to those who fail to meet the conditions. In Baird, McIntosh, and Özler (2011) we show how the protective effects of UCTs relative to CCTs are expressed most strongly immediately subsequent to dropping out of school, a moment at which adolescent girls are particularly susceptible to financial pressures. Many of these individuals come from vulnerable households in need of income support from the government.

Excitement surrounding encouraging findings on asset accumulation from two recent studies – the short-term evaluation of cash transfers given to poor households unconditionally by GiveDirectly, a charitable organization that targets households in rural Kenya by the type of roof material above their house and channels large, lump-sum cash payments to these households unconditionally (Haushofer and Shapiro 2013); and the longer-term evaluation of large, lump-sum, unsupervised cash transfers to groups of young unemployed individuals in Northern Uganda (Blattman, Fiala, and Martinez 2013) – also contributes to the renewed enthusiasm for UCTs.[5] While these programs don't resemble typical cash transfer programs run by governments – conditional or unconditional – their findings are indicative of the fact that, in some settings, credit constraints may be the main obstacle to investment and a sustainable path out of poverty.

Furthermore, some argue that strict conditions could be replaced with gentler 'nudges' to achieve the desired behavior change. Another recent study that has directly experimented with conditions is Benhassine et al. (2013), which evaluates the Tayssir cash transfer program in Morocco. The authors describe one arm as a CCT, where explicit conditions for school attendance were announced and monitored, and the other arm as a 'labeled cash transfer' (LCT). It is 'labeled' because the program was run by the Ministry of Education, where parents of young children had to enroll in the program at the local school and where the headmaster registered and enrolled the children in school while registering them in the program, however, attendance was not monitored: it was made salient to households that the funds were coming from the Ministry of Education under a pro-education program that was trying to raise primary school completion rates in rural Morocco. The authors find that in this context CCTs and LCTs were equally effective in reducing dropouts. This experiment suggests an alternative to the dichotomy between CCT and UCT programs – a hybrid that, in some circumstances, might work better than either.

[4] For a review of the literature on the relative effects of CCT vs. UCT programs, please see McIntosh, Baird, and Özler (2011). Akresh, DeWalque and Kazianga (2013) ran an experiment in Burkino Faso very similar to the Malawi program analyzed here and find that the CCTs and UCTs have similar schooling effects on average, but that CCTs outperformed UCTs for 'marginal children,' such as girls and lower ability children. Baird et al. (2013) provides a systematic review of the schooling effects of CCT and UCT programs.

[5] Interestingly, despite the large size of the transfers, Haushofer and Shapiro (2013) find no effects on health or education outcomes.

It is important to note that, apart from a few recent exceptions, evaluations of cash transfer programs have two shortcomings. First, they tend to focus on a narrow set of outcomes, such as schooling or work. Second, the time horizons of the impact evaluations are relatively short, usually around one- to two-years after baseline.[6] The study presented here tries to fill some of this gap by presenting impacts on a broad range of outcomes of interest for young women five years after baseline or approximately two years after the end of the intervention.

2.2. Study Setting

Malawi, the setting for this research project, is a country of more than 15 million people in southern Africa, more than 80% of whom live in rural areas in 2009 and most of whom rely on subsistence farming. Malawi's 2008 GNI per capita figure of $760 (PPP, current international $) is less than 40 percent of the sub-Saharan African average of $1,973 (World Bank 2010). According to the same data source, net secondary school enrollment is very low at 24 percent. Malawi also has the ninth-highest HIV prevalence in the world with 10.8 percent of adults aged 15-49 infected (UNAIDS, 2013). At 4.2%, HIV prevalence among females aged 15-19 is not only high but also more than *three* times higher than the prevalence of 1.3% among males of the same age group (Government of Malawi 2012).

Within Malawi, Zomba district in the Southern region was chosen as the site for this study for several reasons. First, it has a large enough population within a small enough geographic area rendering field work logistics easier and keeping transport costs lower. Zomba is a highly populated district, but distances from the district capital (Zomba Town) are relatively small. Second, characteristic of Southern Malawi, Zomba has a high rate of school dropouts, low educational attainment, and high HIV rates. At 24.6%, HIV prevalence among women aged 15-49 in Zomba was the highest rate in the country (NSO 2005). According to the Malawi Integrated Household Survey-2 (2004/2005), the biggest reason for dropout from school reported by households was financial. Hence, cash transfers were expected to have sizeable effects in Zomba, where many school-aged girls are at risk of dropping out of school early, becoming pregnant, and marrying early.

2.3. Sampling[7]

---

[6] There is a small but growing body of evidence on the medium/long term effects of cash transfer programs. In the short run, most evaluations of CCT programs focus on outcomes directly related to the condition. For example, in the case of schooling CCTs, they focus on impacts on schooling and related outcomes. In the longer term, evaluations are more likely to look at a wider set of outcomes. For example, Behrman, Parker and Todd (2011) examines the longer term impacts of PROGRESA/Oportunidades in Mexico by exploiting differential exposure to the program. The authors find significant impacts on schooling, reductions in work for younger youth, increases in work for older girls, and shifts from agricultural to non-agricultural employment. Similarly, Barham, Macours, and Maluccio (2013) examines the long-term gains in grade attainment and achievement in math and language among boys 10 years after the start of Nicaragua's RPS CCT program.

[7] For more details on the sample design see Appendix 7.1.

Zomba District is divided into 550 enumeration areas (EAs), which are defined by the National Statistical Office of Malawi and contain an average of 250 households spanning several villages. Fifty of these EAs lie in Zomba city, while the rest are in seven traditional authorities. Prior to the start of the experiment, 176 EAs were selected from three different strata: Zomba city (urban, 29 EAs), near rural (within a 16 KM radius of Zomba city, 119 EAs), and far rural (28 EAs). The choice of a 16 KM radius around Zomba city was arbitrary and based mainly on a consideration of transport costs.

In these 176 EAs, each dwelling was visited to obtain a full listing of never-married females, aged 13-22.[8] The target population was then divided into two main groups: those who were out of school at baseline (*baseline dropouts*) and those who were in school at baseline (*baseline schoolgirls*). The group of *baseline schoolgirls* accounts for more than 85% of the target population within the study EAs. In each EA, all *baseline dropouts* and a percentage of *baseline schoolgirls* were randomly selected for the study. The sampling percentages for baseline schoolgirls differed by strata and age group and varied between 14% and 45% in urban areas and 70% to 100% in rural areas. This procedure led to a total sample size of 3,796 core respondents, of whom 2,907 were baseline schoolgirls and 889 were baseline dropouts.[9]

2.4. Study Design and Intervention

The Zomba Cash Transfer Program is a randomized cash transfer intervention targeting young women in Malawi that provides incentives (in the form of school fees and cash transfers) to current schoolgirls and young women who have recently dropped out of school to stay in or return to school. There are two treatment arms: conditional cash transfers (CCT) and unconditional cash transfers (UCT). This sub-section describes these study arms in detail, including a description of other cash transfer design parameters that were randomized – such as transfer amount and identity of the transfer recipient within eligible households.

Treatment status was assigned at the EA level and the sample of 176 EAs was randomly divided into two groups of equal size: treatment and control. The sample of 88 treatment EAs was further divided into two arms based on the treatment status of *baseline schoolgirls*: (i) CCT arm (46 EAs), and (ii) UCT arm (27 EAs). In the remaining fifteen treatment EAs, no baseline schoolgirls were made offers to receive cash transfers.[10] All

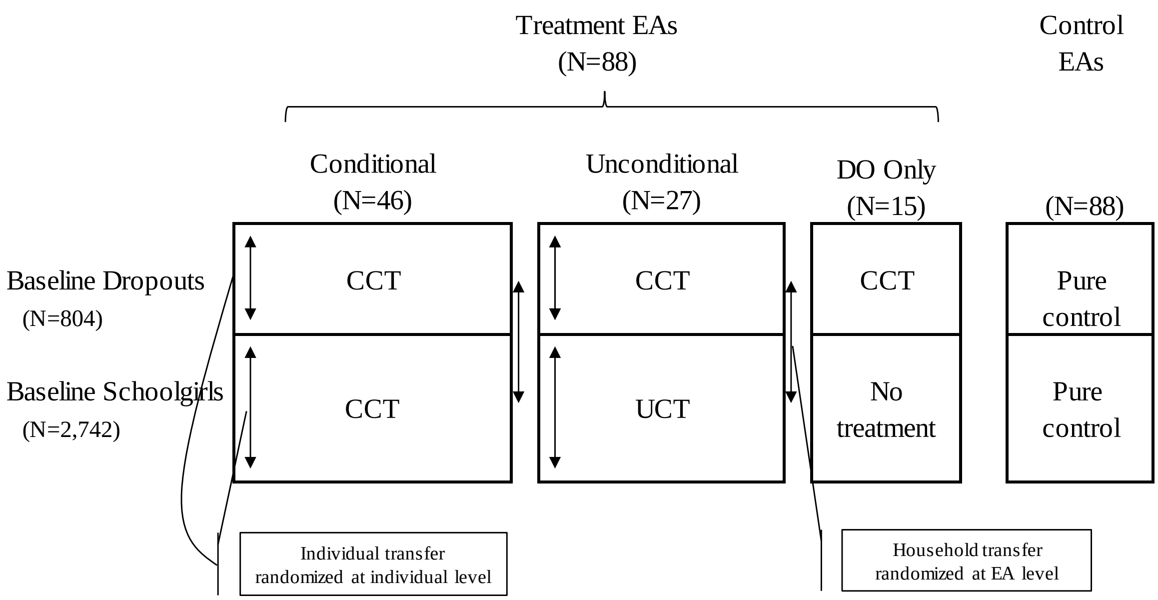
[8] The target population of 13-22 year-old, never-married females was selected for a variety of reasons. The age range was selected so that the study population was school-aged and had a reasonable chance of being or becoming sexually active during the study period. Finally, a decision was made to not make any offers to girls who were (or had previously been) married, because marriage and schooling are practically mutually exclusive in Malawi – at least for females in our study district.

[9] See Appendix 7.2 for a discussion of power calculations conducted for Rounds 2 and 3.

[10] To measure potential spillover effects of the program, a randomly selected percentage (33%, 66%, or 100%) of baseline schoolgirls in each treatment EA were randomly selected to participate in the cash transfer program. In the 15 treatment EAs, where no baseline schoolgirls were offered cash transfers, this percentage was equal to zero. In these 15 EAs, the only spillovers on *baseline schoolgirls* would be from the *baseline dropouts* who were offered CCTs. We do not utilize this random variation in treatment intensity in this paper. Furthermore, a thorough examination of data from the two-year follow-up found no significant spillover effects (Bohren et al. 2014). As a result, the sample of 623 untreated individuals in

baseline dropouts residing in the 88 treatment EAs received CCT offers regardless of the assigned treatment status of *baseline schoolgirls* in their EAs.[11] Figure 1 presents an illustration of the study design. No EA in the sample had a similar cash transfer program before or during the study.

Figure 1: Malawi SIHR Research Design



2.4.1. CCT arm

After the random selection of EAs and individuals into the treatment group, the local NGO retained to implement the cash transfers held meetings in each treatment EA between December 2007 and early January 2008 to invite the selected individuals to participate in the program.[12] At these meetings, the program beneficiary and her parents/guardians were made an offer that specified the monthly transfer amounts being offered to the beneficiary and to her parents, the condition to regularly attend school, and the duration of the program. An example of the CCT offer letters can be seen in Appendix 7.3. It was possible for more than one eligible girl from a household to participate in the program. Transfer amounts to the parents were varied randomly across EAs between $4, $6, $8, and $10 per month, so that each parent within an EA received the same offer. Within each EA, a lottery was held to determine the transfer amount to the young female program beneficiaries, which was equal

treatment clusters were not interviewed at the five-year follow-up and their outcomes are not discussed here.

[11] The treatment arm that experimentally tests the impact of the *conditionality* was applied only in the stratum with baseline schoolgirls and not among the baseline dropouts. The main reason was that, given the small number of baseline dropouts who were eligible for the program, splitting the baseline dropouts into conditional and unconditional treatment groups would have low power to precisely identify treatment effects.

[12] The cash transfer program was implemented by Invest in Knowledge Initiative in 2008 and by Wadonda Consult in 2009. Both implementers were instructed to follow the same set of procedures and, to the best of our knowledge, there were no significant changes in implementation between the two years.

to $1, $2, $3, $4, or $5 per month.[13] The fact that the lottery was held publicly ensured that the process was transparent and helped the beneficiaries to view the offers they received as fair. In addition, the offer sheet for CCT recipients eligible to attend secondary school stated that their school fees would be paid in full directly to the school.[14]

Monthly school attendance for all girls in the CCT arm was checked and payment for the following month was withheld for any student whose attendance was below 80% of the number of days school was in session for the previous month. However, participants were never administratively removed from the program for failing to meet the monthly 80% attendance rate, meaning that if they subsequently had satisfactory attendance, then their payments would resume. Offers to everyone, identical to the previous one she received and regardless of her schooling status during the first year of the program in 2008, were renewed between December 2008 and January 2009 for the second and final year of the intervention, which ended at the end of 2009.

2.4.2. UCT arm

In the UCT EAs, the offers were identical with one crucial difference: there was no requirement to attend school to receive the monthly cash transfers.[15] An example of the UCT offer letters can also be seen in Appendix 7.3. Other design aspects of the intervention were kept identical so as to be able to isolate the effect of imposing a schooling conditionality on primary outcomes of interest. For households with girls eligible to attend secondary schools at baseline, the total transfer amount was adjusted upwards by an amount equal to the average annual secondary school fees paid in the conditional treatment arm.[16] This additional amount ensured that the average transfer amounts offered in the CCT and UCT arms were identical and the only difference between the two groups was the "conditionality" of the transfers on satisfactory school attendance. Attendance was never checked for recipients in the UCT arm and they received their payments by simply presenting at the transfer locations each month.

The UCT experiment was conducted alongside the CCT experiment in the same district. In order to interpret the differential impacts between the two treatment arms, it is important to know what was understood by those in the UCT arm as to the nature of their transfers and to understand the context in which the cash transfer experiment was conducted. Evidence from in-depth qualitative interviews conducted soon after the two-year follow-up indicates that the UCT experiment did not happen in a vacuum. While the rules of the program were

[13] The average total transfer to the household of $10/month for 10 months per year is nearly 10% of the average household consumption expenditure of $965 in Malawi (calculated using final consumption expenditure for 2009, World Development Indicators 2010). This falls in the range of cash transfers as a share of household consumption (or income) in other countries with similar CCT programs. Furthermore, Malawi itself had a Social Cash Transfer Scheme (SCTC), which transferred $12/month plus bonuses for school-age children during its pilot phase (Miller, Tsoka and Reichert 2008). The SCTC did not cover Zomba district during the implementation of Zomba Cash Transfer Program in 2008 and 2009.

[14] Primary schools are free in Malawi, but student have to pay non-negligible school fees at the secondary level. The program paid these school fees for students in the conditional treatment arm upon confirmation of enrollment for each term. Private secondary school fees were also paid up to a maximum equal to the average school fee for public secondary schools in the study sample.

[15] The reader should note, again, that all baseline dropouts in a treatment EA, regardless of the treatment status of baseline schoolgirls in that EA, were offered CCTs.

[16] For details, please see footnote 20 (page 1719) in Baird, McIntosh, and Ozler (2011).

well understood by the girls in the UCT arm, i.e. that UCT girls knew that nothing was required of them to participate in the program and receive their monthly transfers, they were also very much aware of the CCT intervention through friends and acquaintances. Therefore, the UCT intervention took place under a rubric of education that naturally led the beneficiaries to believe that the program aimed to support girls to further their education. The differential impacts of the UCT and CCT interventions should be interpreted in this context.[17]

2.5. Data Sources, Outcomes, and Pre-Analysis Plan

2.5.1. Data Sources

This report presents evidence on the impacts of CCT and UCT interventions separately for baseline dropouts and baseline schoolgirls at one, two, and five years after baseline data collection. Table 1 provides a time line of the program implementation and data collection. Discussions with the Ministry of Education and Vocational Training, as well as meetings with District officers for health, education, and planning were held to assure that the program received acceptance at the national and local levels and the study design was locally appropriate. Additionally, village meetings with traditional authorities, group village heads, and village heads were conducted. All participants provided written informed consent. Additional consent was obtained from parents or legal guardians of all unmarried girls under the age of 18. The study design was approved by ethical review committees at the National Health Sciences Research Council (Malawi, Protocol #569) the University of California at San Diego (USA, Protocol #090378), and George Washington University (USA, Protocol #061037).

Table 1: Intervention and Data Collection Timeline

	Start Date	End Date
Intervention		
Cash Transfer Intervention	February 2008	December 2009
Household Survey		
Baseline Household Survey Data Collection	October 2007	February 2008
Round 2 Household Survey Data Collection	October 2008	February 2009
Round 3 Household Survey Data Collection	January 2010	July 2010
Round 4 Household Survey Data Collection	March 2012	February 2013

Notes: The Round 3 Household Survey also included separate educational tests. The Round 4 Household Survey also included competencies, husband surveys, and child surveys and early childhood development tests. Biomarker data collection also took place during Round 2, 3 and 4.

At baseline (Round 1), following the listing exercise described in Section 2.3, household surveys were conducted, which were revised and repeated at every follow-up round thereafter. At one-year follow-up (Round 2), we conducted a *school survey* that involved visiting every school attended by any of the core respondents (according to self-reported data from the household survey) in 2008; as well as biomarker data collection for HIV, HSV-2, and syphilis in a randomly selected group of 104 EAs (27 UCT, 25 CCT, and 52 control). At two-year follow-up (Round 3), in addition to household surveys, school surveys, and biomarker

[17] For details, please refer to Baird, McIntosh, and Ozler (2011), pages 1719-21.

data collection for HIV and HSV-2, we also developed *mathematics and English reading comprehension tests* and administered them to all study participants at their homes to measure program impacts on learning.[18] During this round, *structured in-depth interviews* were conducted with a small sample of study participants, their parents or guardians, community leaders, program managers, and schools. Finally, at five-year follow-up (Round 4), household surveys included modules for husbands and children of core respondents who were married and/or had children as well as a module to measure basic competencies (described below). Biomarker data collection during R4 included HIV and hemoglobin for the core respondents, HIV for husbands, and anthropometrics for children under the age of five. Details for the various data collection instruments for Rounds 1-3 can be found in Baird, McIntosh, and Özler (2011), Baird et al. (2012), and Baird, de Hoop, and Özler (2013), while the details for Round 4 instruments are described below.

2.5.1.1. Household Survey

The annual *household survey* consisted of a multi-topic questionnaire administered to the households in which the sampled respondents resided. It consisted of two parts: one that was administered to the head of the household and the other administered to the core respondent, i.e. the sampled girl from our target population. The former collected information on the household roster, dwelling characteristics, household assets and durables, shocks and consumption. The survey administered to the core respondent provides detailed information about her family background, schooling status, health, dating patterns, sexual behavior, fertility, and marriage.

In Round 4 the household survey also consisted of a set of questions to try and measure some basic competencies of the core respondent. These competencies included reading and following instructions to apply fertilizer; making correct change during a hypothetical transaction; sending text messages and using a calculator on a mobile phone, and calculating profits for a hypothetical business scenario. They were designed to replace the achievements tests utilized in Round 3, and serve as a measure of a more practical set of skills that might be influenced by increased schooling and needed in the labor market.

In addition to modules administered to the core respondent (and to her parents/guardian if she still lived with them), the Round 4 survey included a module that was administered to the husbands of married core respondents. During this round, early childhood development (ECD) tests were administered for all 3-4 year-old children of the core respondent. These tests consisted of the Malawi Development Assessment Tool (MDAT) for fine motor skills, language, and hearing, which were administered directly to the child and the Strengths and Difficulties Test (SDQ) (http://www.sdqinfo.com/a0.html) which was administered to the core respondent or the guardian responsible for the child.[19]

2.5.2. Outcomes and Pre-Analysis Plan

Prior to the analysis of data from Round 4, a pre-analysis plan was drafted that described in detail the empirical analysis plan and specified the primary and secondary outcomes to be examined. This document, which was designed to help us avoid data mining

[18] In addition, to measure cognitive skills, we utilized a version of Raven's Colored Progressive Matrices that was used in the Indonesia Family Life Survey (IFLS-2).

[19] As this report focuses only on the outcomes of the core respondent, we do not analyze data from the modules concerning outcomes for husbands or children. Future work will include analysis of these data. Future work will also analyze biomarker data on HIV and anemia.

and ad hoc subgroup analysis, was registered in the AEA RCT Registry on August 6, 2013 (AEARCTR-0000036; https://www.socialscienceregistry.org/trials/36). The components of the pre-analysis plan relevant to this report are attached as Appendix 7.4. We discuss the primary and secondary outcomes examined in this report below. The pre-analysis plan provides more detail on exactly how these outcome variables were constructed.

Education and competencies: The primary outcome we examine for educational achievement is highest grade completed. Secondary outcomes reported are highest qualification obtained, as well as whether the core respondent obtained a Primary School Leaving Certificate (PSLC) or Junior Certificate of Education (JCE). As school surveys were not possible in Round 4, these outcomes are based on self-reports.[20] For Round 3, we also report impacts on achievement test scores in mathematics, English reading comprehension, and cognitive ability, all of which were replaced in Round 4 with basic *competencies*, discussed briefly in Section 2.5.1.1 and described in detail in the appendix of the pre-analysis plan (Appendix 7.4) with the instrument shown in full in Appendix 7.5.

Marriage and Fertility: We report program impacts on ever married, ever pregnant, age at first live birth, total live births (primary outcomes); age of first marriage and desired fertility (secondary outcomes).

Health and Nutrition: We examine psychological wellbeing (Rounds 2-4), and the number of meals last week that contained, meat, fish, or eggs during the past week (Rounds 2-4) as secondary outcomes.

Sexual behavior: All the outcomes we report on sexual behavior are defined as secondary outcomes in the pre-analysis plan. They include ever had sex, number of lifetime sexual partners, and being sexually active during the past 12 months (measured on the extensive margin, i.e. for everyone in the study sample); age at first sex, and condom use during the most recent sexual intercourse prior to the survey (measured on the intensive margin, i.e. among those sexually active during the past 12 months). We report program impacts on these outcomes for all three follow-up rounds.

3. Program Implementation

3.1. Program Implementation Challenges

There were a number of challenges faced throughout the implementation of the program, both during the initial set up of the intervention back in 2007, as well as related to the current data collection efforts funded through 3ie.

3.1.1. Challenges at Baseline

[20] On the possibility of differential bias in self-reported measures of school enrollment and attendance, please see Baird and Özler (2012).

At the onset of implementing the SIHR study we encountered problems associated with villagers' fears of what are commonly called 'blood suckers' in the southern part of rural Malawi. Seemingly synonymous with what may be called a vampire, periodically, fears of these 'blood suckers' surface in rural Southern Malawi, especially in Chiradzulu, the district originally selected to be the site for this study. The reports are never consistent – there is never any concrete evidence of the existence of such beings (it is debated whether they are humans or spirits) or what it is that they actually do to suck their victims' bloods or what they do with the blood, but the consequences of such flare-up of fears is substantial: villagers set up road blocks after dark, establish curfews, sleep outside of their houses around a fire together (it is rumored that the 'blood suckers' can come in through the cracks in the walls or the ceiling of the house), and protect themselves from anyone, especially strangers, suspected to be a blood sucker using weapons.

While we were in the field in September, 2007, an ambulance was attacked and its driver injured as he was driving to a village at night to pick up a patient. In a separate incident, a police car was attacked by a group of villagers. Any stranger going through a village is suspect, especially at night, and even local officials, such as village headmen or traditional authorities are not spared suspicion: often, villagers think that these authorities have been bribed by the 'blood suckers' to ease their access to the village and to deny the existence of 'blood suckers'. The result is an atmosphere of distrust of any stranger (defined as an individual not living in the village), sleepless nights, and sometimes violence resulting in severe injuries to people and significant damage to property.

Unfortunately, the blood sucker rumors started flaring up again before our field work preparations began in September, 2007 in Chiradzulu. After consulting with our survey firm, our local counterparts, local officials, and experienced field workers, we were assured that the blood sucker fears would not affect our field work if we held meetings with the local officials and explained our purposes for being there and interviewing people. Hence, after holding such meetings and obtaining letters of approval from local officials (including the District Commissioner – the highest ranking government official in a district in Malawi), we started our field work on October 5, 2007 in seven randomly chosen EAs in Chiradzulu.

The first couple of days went quietly. There were some refusals to be interviewed here and there, and the teams were not able to work after dark due to road blocks, but otherwise things seemed peaceful. On the third day, however, we had two teams come back from the field with reports of having been chased away by villagers because of fears of 'blood suckers'. As one of the teams had their vehicle's front window shield smashed, and the other team was chased away (in the presence of the village headman) by panga knives and stones, we have held an emergency meeting and considered our options. After a short discussion, it was clear that the safety of the field workers could not be ensured in Chiradzulu, so we decided to abandon it for a neighboring district.

To continue field work, we needed to find a district that had similarly high rates of school dropout and HIV/AIDS among young women. In addition, the district had to be a reasonable size to keep transport and survey costs to a minimum. We had to make sure that it would be a district to which 'blood sucker' problems would spread with only a low likelihood, as well as making sure that a poverty-targeted UNICEF cash transfer program was not being implemented (and would not be over the next two years) in that district. After a few days of analyzing data and discussions with our local counterparts and our survey firm, we decided that Zomba, with its high HIV/AIDS rates, and a good mixture of urban and rural areas, was the best choice. Zomba town is the former capital of Malawi, and the district within which it lays borders Chiradzulu. In many ways, rural Chiradzulu and rural Zomba resemble each other and people migrate from both areas to the nearby cities of Blantyre and Zomba for education and work.

Even though Zomba district is not very large, it still is much bigger than Chiradzulu in size, with its EAs more spread out. To keep transport costs as well as interview and listing times to a minimum, we stratified Zomba into three areas: Zomba town, rural Zomba within the 16 KM radius of Zomba town, and rural Zomba outside the 16 KM radius of Zomba town. After randomly sampling new EAs from these three strata, the field teams went back to work.

After completing work in about 42 EAs in October 2007, with infrequent problems with 'blood sucker' fears, similar problems flared up in one traditional authority (TA) where the field teams were conducting interviews. It seemed that the problem was spreading throughout the Southern Region of Malawi and even Zomba district was not spared. After discussions with the field teams, we decided to develop a formal 'sensitization process', where the field teams would spend a longer time in each village to explain the intervention and the study and would not go into villages to start the listing exercise until they had had a village meeting and/or the group village headmen had a chance to inform most of the villagers. This added significant amount of time to the field work (as sensitization took additional time each time the teams moved to a new area in Zomba) and added a wrinkle to the research design (as from that point onwards, the people being interviewed would know that there is a possibility of receiving social assistance, whereas before they only knew about a research study about young people and schooling).

Sensitization was a success and even though time-consuming, well-worth the effort. Without it, the whole project would have been in jeopardy as we may well have had to stop baseline data collection.

3.1.2. Challenges during Round 4 Fieldwork

The challenges during the most recent round of fieldwork were not nearly as severe and were largely due to fuel shortages. Specifically, frequent fuel shortages meant that fieldwork would often be delayed due to lack of fuel or vehicles would have to wait in long lines early in the morning to refill, thus wasting a large proportion of the day. In addition, given that this fieldwork is now 5 years after baseline, many core respondents were harder to track, thus further increasing the total time for fieldwork and substantially increasing our costs. In order to stay within budget, the research team decided not to re-interview the control girls in treatment EAs, a decision that was based on the fact that the research team found minimal short-run spillover effects. With this minor deviation from our initial protocol, the research team was able to do the remainder of the work within budget.

4. Impact Findings

The evaluation of the impact of the ZCTP utilizes the experimental design of the intervention. The experimental study design gives us a reliable source of identification. To estimate intention-to-treat effects of the program in each treatment arm on our primary outcomes we employ a simple reduced-form linear probability model of the following type:

(1) $$Y_i = T_i^C \gamma^C + T_i^U \gamma^U + X_i \beta + \varepsilon_i,$$

where Y_i is an outcome variable for individual i, T_i^C and T_i^U are binary indicators for offers to be in the CCT and the UCT arms, respectively, and X_i is a vector of baseline characteristics.

Note that for baseline dropouts we only have the CCT binary indicator. The standard errors ε_i are clustered at the EA level which accounts both for the design effect of our EA-level treatment and for the heteroskedasticity inherent in the linear probability model.

In all regressions, we include baseline values of the following variables as controls: a household asset index, highest grade attended, a dummy variable for having started sexual activity, and dummy variables for age. These variables were chosen because they are strongly predictive of schooling outcomes and, as a result, improve the precision of the impact estimates. We also include indicators for the strata used to perform block randomization – Zomba Town, within sixteen kilometers of the town, and beyond sixteen kilometers (Bruhn and McKenzie 2008). Age- and stratum-specific sampling weights are used to make the results representative of the target population in the study area.

The full set of tables on the impact results are presented in Appendix 7.6. In this section we focus on a summary of some of the key results. We first provide a look at some descriptive statistics and attrition. We first look at baseline schoolgirls and then at baseline dropouts. The baseline schoolgirl analysis largely focuses on differential impacts between the CCT and UCT arm, while the analysis of baseline dropouts, where the UCT experiment was not conducted, focuses on the CCT arm versus the control group. We look at results for both the core respondent and her children.

4.1. Attrition and Descriptive Statistics

Before turning to a discussion of attrition and some basic descriptive statistics of our sample, we first want to remind our reader of the two sub-groups we analyze separately throughout our analysis: *baseline schoolgirls* and *baseline dropouts*. There are a number of reasons for this separation. First, the schooling condition works differently on these two groups – for *baseline dropouts* it brings them back into school, while for *baseline schoolgirls* it prevents them from dropping out. Second, as described in Section 4.1.1 below, these groups look vastly different across a host of baseline characteristics and thus are best viewed as separate populations. Finally, baseline dropouts are a group that is often ignored in the analysis of CCT programs even though the size of this population in many settings in not negligible. Thus, we feel that providing results separately for this group may provide some important insights.

Table 2 provides the attrition analysis. Column one looks at the sample of baseline schoolgirls that were re-surveyed in both Round 3 and Round 4. First, the mean in the control group is 0.864 indicating the share of baseline schoolgirls in the control group who were resurveyed in Round 3 and Round 4, implying an attrition rate of less than 15%. There are no significant differences between the CCT and UCT treatment arms; however, both treatment groups were more likely to be part of the sample than the control group (5.4% and 4.6% respectively). Thus, any comparison of outcomes in either treatment arm vs. the control group among baseline schoolgirls must be interpreted with caution. Comparisons of CCT vs. UCT do not face such issues. Turning to baseline dropouts, column (2) of Table 2 show no differential attrition issues for this subgroup.

Table 3 presents descriptive statistics for baseline schoolgirls and baseline dropouts and also shows baseline balance for the sample of young women who were resurveyed in Round 3 and Round 4. Out of 32 coefficients, five are significantly different than each other, four of which are related to highest grade and age. To account for these baseline differences all regressions control for highest grade, age and asset index. Also note that all girls were never married at baseline. Table 3 further highlights some important differences between baseline schoolgirls and baseline dropouts to further motivate our decision to analyze these two groups separately. Most notably, 2.1% of baseline schoolgirls in the control group had ever been pregnant at baseline compared to 44.7% among baseline dropouts. This number is further supported by the percentage of core respondents reporting never having had sex (80.3% vs. 30.5%). Moreover, baseline dropouts are older, have lower levels of education, are poorer, and are less likely to come from an urban area.

4.2. Baseline Schoolgirls

The short term results of the evaluation of SIHR found large and significant impacts on schooling, learning, and marriage and fertility, with some important differences between the CCT arm and the UCT arm. The five year analysis presented here both looks at whether these results were sustained two years after the program ended.

4.2.1. Education and Competencies

Table 4A presents program impacts among *baseline schoolgirls* on highest grade completed and highest qualification obtained. The estimates suggest little, if any effects, on these outcomes for either treatment group: there are some modest and statistically significant effects in Round 3, i.e. immediately at the end of the two-year intervention for the CCT arm, but these effects get smaller and become statistically insignificant by Round 4. Highest grade completed has increased by about 0.1 years in both groups, with no difference between them for any outcome in any round statistically significant. The reader should note that the mean number of years completed in the control group in Round 4 is 10.4 and that 88% of the control group has passed the PSLC exam, meaning that they successfully completed primary school. These figures indicate that close to 90% of cash transfer in this stratum were inframarginal for primary school completion. The potential effect of cash transfers on these outcomes is limited by definition and targeting students at risk of dropping out may prove to be more cost-effective.

Table 5A presents program impacts on test scores in Round 3 and basic competencies in Round 4. There were short-term impacts on cognitive ability, mathematics, and English reading comprehension in the CCT group (with the last effect being significantly higher than that in the UCT arm), but these effects did not translate to better performance in practical competencies measured at the five-year follow-up. The index of competencies increased in both the CCT and UCT groups by less than 0.1 standard deviations (SD) and only one out of 10 coefficient estimates is significant at the 10% level (UCT group is more proficient in sending text messages successfully). The results suggest that the earlier improvements in learning were too small to make a difference in the longer-run; that learning decays quickly; or that improved test scores do not translate to more practical life skills, such as making correct change during a market transaction, reading and following instructions, or calculating profits.[21]

4.2.2. Marriage, Fertility, and Desired Fertility

Table 6A presents program impacts related to marriage, actual fertility, and desired fertility. Focusing first on the CCT group of *baseline schoolgirls*, we find no effects on any of the six outcomes presented in this table. By Round 4, 40% of the *baseline schoolgirls* in the control group were ever married, 50% ever pregnant, and with the average age of marriage (among those married) being 18.6. There are no changes for these variables in the CCT arm, although the trend from Round 3 to Round 4 is in the expected direction. Interestingly, desired fertility is also unchanged in this group – with the mean number of children desired approximately three.

Switching to examining impacts in the UCT arm, the coefficient estimates confirm earlier findings that UCTs were effective in reducing marriage and pregnancy rates among *baseline schoolgirls* during and immediately after the program. However, we see an almost complete reversal of these outcomes at the five-year follow-up: there are no longer any differences in ever married, ever pregnant, or the total number of live births between the UCT group and either of the two other study arms. We find that the age at first marriage increased by half a year by Round 4, which is consistent with the fact that girls in the UCT arm who delayed marriage were quickly married following the end of the intervention. The increase of 0.13 years in the age at first live birth is not statistically significant. As with the CCT group, desired fertility also remains unchanged in this group.

These findings suggest that the UCT effects were temporary and due to an income effect. Given that the significant and sizeable effects on marriage and pregnancy during the program disappeared quickly and that there are no effects on desired fertility, it's hard to reach any conclusion other than cash transfers having had no sustained effect on fertility.

4.2.3. Health and Nutrition

When it comes to health and nutrition outcomes, the picture is one of significant and meaningful effects during the program disappearing by Round 4. Looking at nutritional intake, we find that both types of cash transfers led to a significant increase in the number of meals during which a source of protein (meat, fish, or eggs) was consumed during Rounds 2 and 3: the effects are on the order of half a meal over a control mean of approximately four such meals per week, or about a 10% improvement. These effects disappeared two years after the cash transfers were stopped.

[21] The authors developed the competencies and the reader might question the relevance of these 'skills.' However, both piloting before Round 4 and current analysis suggests that these skills are correlated with previous test scores and highest grade completed. It seems that the increases in test scores (also around 0.1 standard deviations and lowest for math skills) at the end of the program may have been too small to cause improved performance at the five-year follow-up.

Second, examining the prevalence of psychological distress (using a binary indicator based on the GHQ-12 index), we confirm earlier findings (Baird, de Hoop, and Özler 2013) of large improvements in mental wellbeing during the program that disappeared as soon as the cash transfers stopped. The analysis here confirms that there were no further changes between Rounds 3 and 4. A small, but reliable source of income seems to reduce psychological distress (i.e. the possibility of suffering from mild anxiety and depressive disorders), but does not cause sustained improvements beyond the intervention period.

4.2.4. Sexual Behavior

Tables 8A and 9A shows that neither type of cash transfer program had any effect on sexual behavior – either on the extensive margin (such as onset of sexual activity or number of partners) or on the intensive margin (age at first sex, condom use, age of partner). By Round 4, close to 70% of baseline schoolgirls were sexually active with an average of approximately one sexual partner.

4.3. Baseline Dropouts

4.3.1. Education and Competencies

Table 4B presents program impacts among *baseline dropouts* on highest grade completed and highest qualification obtained. We can see that, unlike the effects among *baseline schoolgirls*, the effects in this group of 'higher risk' girls are much larger and much more durable over the long term.[22] For example, the program caused an increase of 0.6 years of schooling completed over a control mean of 7 years. Given that primary school is 8 years in Malawi, this improvement caused an 8 percentage point increase in passing the PSLC, over a control mean of 37%, i.e. a 21% increase in primary school completion rates. Unlike baseline schoolgirls, the gains obtained during the program in schooling years and qualifications do not disappear at the five-year follow-up because the control group is unlikely to return to school. The program affected junior secondary school completion quite strongly as of the completion of the treatment (~5 pp increase over a base of 12%), and that while the gap in this variable is no longer significant by Round 4 it remains ~3 pp, and so has not seen the sharp 'bounce-back' that characterizes the post-treatment educational responses in the baseline schoolgirl group. The large effects of cash transfers in this group highlight the potential importance of trying to find ways to target 'at risk' children for schooling interventions.[23]

[22] This particular sub-group is often ignored in school-based programs (because they are already out of school), but may be a particularly vulnerable group. In our study sample, they were, on average, older, more likely to come from poorer households, and much more likely to have ever been pregnant – compared with *baseline schoolgirls*.

[23] The difficulty for policy makers is designing a successful targeting and incentive scheme that brings out-of-school children back into the system, without providing perverse incentives for dropping out of school.

Table 5B presents program impacts on test scores in Round 3 and basic competencies in Round 4. Here the findings are much more similar to the CCT group of baseline schoolgirls presented above: short-term gains in cognitive and math skills at the two-year follow-up did not translate to improved performance on basic competencies in Round 4. This consistent pattern of short-term improvements in test scores combined with no improvement in long-run competencies has two potential explanations. One of these is that the competencies simply failed to measure variation in abilities in a useful way. Arguing against this is the strong observational correlation between test scores and competencies (for example, the correlation between the standardized math score and the standardized total competency for baseline schoolgirls is 0.488 (p<0.000) which implies that a one standard deviation increase in the math test leads to a 0.488 standard deviation increase in the total competency). It therefore appears that a more likely explanation for this is that the modest amount of increased learning engendered by the CCT programs did not amplify the kinds of practical abilities that the competencies were geared to measure. This is in and of itself an interesting result in terms of the specific types of human capital that we can hope to improve through the use of schooling CCT programs.

4.3.2. Marriage, Fertility, and Desired Fertility

As with schooling outcomes, CCTs had large effects on marriage and fertility that were sustained five years after baseline (Table 6B). *Baseline dropouts* were 14.2, 15.7, and 10.3 percentage points less likely to have been ever married at one-, two-, and five-year follow-ups, respectively. The corresponding reductions were 5.8, 8.1, and 3.8 percentage points for being ever pregnant. Furthermore, there was a negative gradient in the total number of live births with reductions of 0.01, 0.095, and 0.152 children during Rounds 2-4, respectively. This means that two years after the cash transfers were discontinued; the number of children in the CCT group was 0.15 lower than the control mean of 1.35, or a 10% reduction that is statistically significant at the 99% level of confidence. Age at first marriage and first birth were similarly higher by 0.41 and 0.31 years, respectively. Finally, there was a modest decline of 0.16 children in desired fertility (from a control mean of 3.22, only significant at the 90% level).

4.3.3. Health and Nutrition

Table 7B suggests that the large gains in schooling attainments and the reductions in marriage and fertility rates were not accompanied by any obvious gains in health outcomes. There were no improvements in mental health even during the program. The lack of income or schooling effects on these outcomes is curious and somewhat surprising and deserves more analysis.[24] There were no effects on meals taken either.

[24] Baird, de Hoop, and Özler (2013) speculates that the lack of improvement in psychological wellbeing may be due to the significant change in lifestyle for *baseline dropouts*, caused by returning to school at a slightly older age.

4.3.4. Sexual Behavior

Finally, Table 8B and 9B suggest that CCTs delayed the onset of sexual activity among *baseline dropouts* – consistent with the effects on schooling and fertility. However, these effects were limited to the period during and immediately after the program: virtually the entire group of *baseline dropouts* (97%), which is, on average, one and a half years older than *baseline schoolgirls*, is sexually active by Round 4. Table 9B shows that CCTs did not cause any changes in condom use or age at first sex among those sexually active.

5. Policy Recommendations

Evaluations concerned with the impact of interventions to improve outcomes for school-aged children or adolescents in developing countries usually measure outcomes within a short time frame, such as one or two years, although longer-term studies are slowly emerging (e.g. Behrman, Parker and Todd 2011; Duflo, Dupas, and Kremer 2012; Barham, Macours, and Maluccio 2013). However, while short-term improvements, such as significant increases in current welfare, are valuable in and of themselves, the purpose of many such interventions is to improve future outcomes, i.e. welfare improvements that last well after the end of the interventions. Cash transfer programs, especially CCTs, explicitly aim to reduce current poverty by providing a safety net for poor families now and future poverty by providing incentives to accumulate human capital (Fiszbein and Schady 2009). As such, it is important to evaluate their longer-term effects to assess whether the short-term improvements translate into sustained gains after the end of such programs.

Such gains require not only behavior change in the short-run, but the accumulation of some sort of capital – physical, human, or social – during and after the program. Without a significant accumulation of some combination of skills, health, knowledge, information, and networks, sustained future gains from programs targeting school-aged children and young women are unlikely.

In this paper, we present a variety of outcomes – educational attainment, learning and skills, marriage and fertility, health and nutrition, and sexual behavior – for a group of school-aged girls in Malawi who were part of a cash transfer experiment. The outcomes we examine were measured after approximately one, two, and five years after baseline data collection. Our findings indicate that while CCTs and UCTs had significant effects on many of these outcomes during and immediately after the program, these effects had mostly dissipated two years after the experiment ended. We argue that the lack of medium-term effects, especially among *baseline schoolgirls*, is due to the fact that there was an insufficient amount of human capital accumulation. In fact, the only large and sustained effects (in marriage, fertility, and desired fertility) were among baseline dropouts, which was accompanied by large increases in school attainment in this more 'at risk' group of adolescent girls and young women, many of whom returned to school in response to offers of CCTs.

Our findings also suggest that researchers analyzing short-term effects should be careful in interpreting them. Short-term findings (two years after baseline) from this study suggested that while CCTs were more successful in improving schooling outcomes than UCTs, UCTs were more successful in reducing teen pregnancies and early marriages. Five years from baseline, these significant effects have disappeared: the outcomes are indistinguishable between the control, CCT, and the UCT groups. The trends in marriage, and fertility in the UCT group between Round 3 and Round 4 (i.e. trends after the cash transfers stopped) are particularly telling: all of the significant gains obtained during the program were wiped out by trends in the other direction. In other words, UCTs simply delayed pregnancies and marriages rather than preventing them: the cash transfer program was akin to pushing a pause button for these school-aged girls for two years, but once that button was released they engaged in these behaviors at rates *higher than* the counterfactual and hence within two years they had completely caught up with the trajectory on which they would have been had the program never been in place. The UCT program does not seem to have enabled them to accumulate physical or human capital that could translate into a different trajectory in life.[25]

While differential post-program trends saw *baseline schoolgirls* in the UCT group converge to the control group, the picture is different among *baseline dropouts*. While the total numbers of live births for the control and CCT groups were very similar at Round 2, girls in the CCT group had approximately 0.1 less births at Round 3 and 0.15 births at R4 – meaning that their fertility was trending in the opposite direction to baseline schoolgirls who received UCTs. Having completed 0.6 more years of schooling and being 8 percentage points more likely to complete primary school, this group of at risk girls had lowered fertility by 10% five years after baseline, a figure that is consistent with their lower reported levels of desired fertility.

A possible takeaway from these findings is not only that human capital accumulation is important to measure in the short-run, but also that any such accumulation has to be substantive enough to translate into improvements in final outcomes later on. *Baseline dropouts*, who were unlikely to return to school on their own (17% reported being in school at Round 2) but did so in large numbers when offered CCTs (61% reported being in school at Round 2), form one group in which causing large changes in educational attainment was possible. Even among this group, while the effect size of 8 percentage points (over a control mean of 37%) on completing primary school would be considered large by most researchers and policymakers, it indicates that the transfers still failed to see many girls through the end of primary school. And, despite these gains in attainments, we do not see any improvements in basic competencies in this group. As Pritchett (2013) states, raising performance is 'wicked hard.'

The complete lack of effects among *baseline schoolgirls* and some sustained effects among *baseline dropouts* also points to the potential importance of targeting these interventions to the right groups. As indicated earlier, by Round 4, 88% of *baseline schoolgirls* had completed primary school, while 54% had obtained a junior secondary school certificate. In such groups, CCTs are inframarginal for a large subset of individuals because they would have attended school without the subsidy anyway. Compared with this group, *baseline*

[25] When we liken the UCT program to a pause button, we do not mean to minimize the welfare improvements caused by the additional income these girls and their families had during the two-year program. For example, our own findings suggest that individuals in the treatment arm had much higher levels of psychological wellbeing during the program, they ate more, were more likely to attend school, etc. Delaying pregnancies may have improved outcomes for the children they conceived later; delaying marriages may have improved the quality of their matches in the marriage market. The effects of the program on the children and marriages of the core respondents is the subject of future study. Our point here is that many of these programs aim to improve future outcomes rather than being palliative. As such, evaluating the effects of such programs in the short run can paint misleading pictures with respect to their overall effects on final outcomes.

dropouts come from much more disadvantaged backgrounds. While schooling status at baseline is not a particularly attractive indicator for targeting, finding cheap and effective targeting mechanisms would significantly improve the cost-effectiveness of these programs.

Our findings may also suggest that the two-year cash transfer program aimed at girls mainly eligible to attend the final years of primary school or higher could have been too short and too late. Many programs target children earlier and last longer. Cash transfers that lasted a few more years could have caused more permanent effects on health, nutrition, and fertility; a program that started earlier could have prevented dropouts and pregnancies before it was too late. While it is not possible for us to extrapolate the effects of our program for a younger cohort, structural modeling combined with experimental data from this study can tell us whether the effects would have been different had the program lasted longer.[26]

Understanding target populations and key outcomes of interest is critical for designing both the right type of policy, as well as the optimal length of the program. In this particular case, there were important differences in the impacts of CCTs and UCTs in the short run, as well as in the impacts of CCTs between *baseline schoolgirls* and *baseline dropouts*. The intervention that will have the largest desired impact will ultimately depend on the outcomes that the policy maker cares most about, as well as the individuals they are hoping to reach. There are likely to be important trade-offs between policy options.

Earlier research from the SIHR study helped to build the current enthusiasm for the surprising ability of unconditional transfer programs to improve a wide range of outcomes in households that are receiving the transfers. We expect this longer-term analysis to temper this enthusiasm, showing these benefits in this case to have been entirely transitory in nature. In this sense the current study re-focuses our attention on the difficult work of building up durable capital stocks, and suggests the particular importance of bolstering human capital among the most vulnerable segments of this adolescent population.

6. References

Akresh, R., de Walque, D. & Kazianga, H. 2013, "Cash Transfers and Child Schooling: Evidence from a Randomized Evaluation of the Role of Conditionality", The World Bank, Policy Research Working Paper No. 6340.

Baird, S., de Hoop, J. and Özler, B. 2013, "Income Shocks and Adolescent Mental Health," *Journal of Human Resources*, vol. 48, no. 2, pp. 370-403.

Baird, S., Garfein, R., McIntosh, C., and Özler, B. 2012, "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial," *The Lancet,* Vol. 379(9823), pp. 1320-1329.

Baird, S., McIntosh, C. & Özler, B. 2011, "Cash or Condition? Evidence from a Cash Transfer Experiment", *Quarterly Journal of Economics,* vol. 126, no. 4, pp. 1709-1753.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M., 2012, "Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda", *Unpublished Manuscript.*

[26] We are currently conducting such an exercise in collaboration with other researchers.

Barham, Tania, Karen Macours, and John A. Maluccio. 2013. "Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions." *American Economic Review*, 103(3): 467-71.

Behrman, J.R., Parker, S.W. & Todd, P.E. 2011, "Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Follow-up of PROGRESA/Oportunidades", *Journal of Human Resources,* vol. 46, no. 1, pp. 93-122.

Benhassine, N., Devoto, F., Duflo, E., Dupas, P. and Poliquen, V. 2013, "Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education", NBER Working Paper No. 19227.

Blattman, C., Fiala, N. & Martinez, S. 2013, "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda", *Quarterly Journal of Economics, Forthcoming.*

Case, A., Hosegood, V. & Lund, F. 2005, "The reach and impact of Child Support Grants: Evidence from KwaZulu-Natal", *Development Southern Africa,* vol. 22, no. 4, pp. 467-482.

Duflo, E., Dupas, P., and Kremer, M. 2012, "Education, HIV, and Early Fertility: Experimental Evidence from Kenya", Working Paper.

Field, E. and Glennerster, R. 2013, "Empowering Girls in Rural Bangladesh", (accessed May 6, 2013), http://www.povertyactionlab.org/evaluation/empowering-girls-rural-bangladesh.

Fiszbein, A. & Schady, N. 2009, *Conditional Cash Transfers: Reducing Present and Future Poverty,* With Francisco H. G. Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto, and Emmanuel Skoufias; Washington, D.C.:World Bank.

Garenne, M. 2008, *Fertility Changes in Sub-Saharan Africa.* DHS Comparative Reports No. 18. Calverton, Maryland, USA: Macro International Inc.

Government of Malawi., 2012, "2012 Global AIDS Response Progress Report: Malawi Country Report for 2010/2011.

Haushofer, J. & Shapiro, J., 2013 "Household Response to Income Changes: Evidence from an Unconditional Cash Transfer Program in Kenya", Working Paper.

Marston M, Slaymaker E, Cremin I, et al. 2009, "Trends in marriage and time spent single in sub-Saharan Africa: a comparative analysis of six population-based cohort studies and nine Demographic and Health Surveys." Sexually Transmitted Infections 85(Suppl. 1), i64–i71.

Miller, C., Tsoka, M., Reichert, K. 2008, *Impact evaluation of the Mchinji Social Cash Transfer*. Report to the Government of Malawi and other stakeholders including UNICEF and USAID. Center for Global Health and Development. Boston, MA.

National Statistical Office (NSO) [Malawi] and ORC Macro. 2005, Malawi Demographic and Health Survey 2004. Calverton, Maryland: NSO and ORC Macro.

Patel, V., Fisher, A.J., Hetrick, S. and McGory, P. 2007, "Mental Health of Young People: a Global Public Health Challenge." *Lancet*, 369, 1302-1313.

Saavedra, J.E. & Garcia, S., 2012, "Impacts of Conditional Cash Transfers on Educational Outcomes in Developing Countries: A Meta-analysis." RAND Corporation Working Papers, WR-921-1.

UNAIDS. 2013, "Global report: UNAIDS report on the global AIDS epidemic 2013," UNAIDS: Geneva. Available from: http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf

United Nations Population Fund (UNFPA). 2012,"Status Report Adolescents and Young People in Sub-Saharan Africa Opportunities and Challenges", South Africa.

World Bank. 2010. World Development Indicators. Washington, DC: World Bank.

World Bank. 2013. World Development Indicators. Washington, DC: World Bank.

7. Appendix

7.1. Sample Design

Zomba District contains 550 EAs defined by the National Statistical Office of Malawi. Each EA contains an average of 250 households spanning several villages. Zomba City includes 50 EAs, while the remaining 500 EAs lie within seven traditional authorities (TAs). Prior to the start of the trial, 176 EAs were selected from three different strata: Zomba City (urban, 29 EAs), near rural (within a 16 KM radius of Zomba City, 119 EAs), and far rural (28 EAs). The choice of a 16 KM radius around Zomba City was arbitrary and was based mainly on a consideration of transport costs. Of the 50 EAs in Zomba City, 21 were excluded per the advice of local experts who deemed these EAs to be too affluent for the proposed intervention. In each of the two rural strata, with the exception of one TA that was unsafe for field work, the study EAs were randomly selected from the universe of all EAs.

In September 2007, after selecting sample EAs, all households were listed in the 176 sample EAs using a short two-stage listing procedure. The first form, Form A, asked each household the following question: 'Are there any never-married girls in this household who are between the ages of 13 and 22?' This form allowed the field teams to quickly identify households with members fitting into our sampling frame, thus significantly reducing the costs of listing. If the answer received on Form A was a 'yes', then Form B was filled to list members of the household to collect data on age, marital status, current schooling status, etc. From this we could categorize the target population into two main groups: those who were out of school at baseline (*baseline dropouts*) and those who were in school at baseline (*baseline schoolgirls*).

These two groups comprise the basis of our sampling frame. Due to their small number (approximately 5 per EA), all eligible baseline dropouts were sampled to participate in the overall study. In the cohort of baseline schoolgirls, a subset of eligible individuals was randomly selected for the study. The sampling percentages for this cohort differed by geographic strata and age-group and varied between 14% and 45% in urban areas and 70% to 100% in rural areas. This sampling procedure yielded a baseline study sample of 4,051 individuals of whom 3,796 (94%) were enrolled and completed a baseline interview at the end of 2007. Of these 3,796 study participants, 889 were baseline dropouts and 2,907 were baseline schoolgirls.

7.2. Power Calculations (for initial intervention)

The random, clustered sample of girls and young women in Zomba was chosen to enable the research team to identify treatment effects on the outcome variables of interest with reasonable confidence. Power calculations indicate that our sample size of 3,796 individuals (in 176 enumeration areas) will allow us to detect moderate treatment effects being significantly different than zero with confidence (90 percent) and considerable power (80 percent). We present power calculations on school enrollment.

As this is a cash transfer program conditional on schooling for one of the treatment arms, it is important that the study be powered to detect not only overall treatment effects on schooling, but also for each of the two treatment arms and for various sub-groups (baseline dropouts and schoolgirls). Table 1.a presents power calculations for detecting one-year treatment impacts of the program across the treatment arms. The figures for the "*observed* probability of enrolment in control", "range of mean enrolment in control EAs", and "*observed* probability of enrolment in treatment" come directly from our analysis of the one-year impact of the program using the follow-up data. "Minimum probability of detectable success in treatment" (in column 4) is the minimum enrolment rate that our power calculations tell us we can detect to be significantly different than control at follow-up. This means that, if a treatment effect can be detected, it has to be outside the range given by columns (2) and (4).

Table 1.b then makes informed *projections* on two-year impacts. As mentioned above, in all the calculations, we use alpha = 0.1 and power = 80%. We utilize the "Optimal Design (OD)" software, which allows us to take into account the fact that our intervention has a randomized, clustered design that is evaluating impacts for continuous or binary outcome variables.

Table A: Power calculations for "school enrolment" (*observed* one-year impact)

(1)	(2)	(3)	(4)	(5)
T/C	*Observed* probability of success in C	Plausible range in C	Minimum detectable probability of success in T	*Observed* probability of success in T
All	70.8	50-90	76.4	80.6
T1/C1	18.6	10-50	26.5	63.2
T2/C2	87.0	75-95	90.9	90.2
T2a/C2	87.0	75-95	91.4	90.3
T2b/C2	87.0	75-95	91.6	90.0
T2a/T2b	90.0*	80-95	95.1⁺	90.3⁺

Note: All figures are in percentages. Alpha = 0.1, power = 0.8. Figures in columns (2), (3), and (5) are based on the impact analysis conducted using the follow-up data.* Probability of success in T2b (i.e. for *unconditional* transfer recipients) + Probability of success in T2a (i.e. for *conditional* transfer recipients)

Table B: Power calculations for “school enrolment” (*projected* two-year impact)

(1)	(2)	(3)	(4)	(5)
T/C	*Projected* probability of success in C	Plausible range in C	Minimum detectable probability of success in T	*Projected* probability of success in T
All	55.0	35-75	60.7	72-79
T1/C1	10.0	5-35	16.5	50-60
T2/C2	70.0	55-85	75.7	80-85
T2a/C2	70.0	55-85	76.4	85-90
T2b/C2	70.0	55-85	76.7	80-90
T2a/T2b	80.0*	70-90	87.2+	85-90+

Note: All figures are in percentages. Alpha = 0.1, power = 0.8. Figures in columns (2), (3), and (5) are projections of two-year impacts based on the one-year impact analysis conducted using follow-up data.

* Probability of success in T2b (i.e. for *unconditional* transfer recipients)

+ Probability of success in T2a (i.e. for *conditional* transfer recipients)

Table A shows that our study is powered to detect meaningful changes in enrolment even after just one year of the program. Relative to the impacts we are observing with the sub-sample of the data currently available to us, the power of the study is quite high for the combined effect of the two treatment arms and for baseline dropouts, and is sufficient to detect the impact from each of the two treatment arms separately for schoolgirls. Table B shows that, as beneficiaries and controls continue to diverge in terms of their schooling status and attainment, the study is likely to detect even larger and statistically significant impacts after two years. Even if these impacts dissipate substantially two years after the program, we are still well powered to identify any economically relevant impact.

Power calculations also show our study has enough power to detect quite small changes in marriage rates (and much larger changes among baseline dropouts) after one year. The impact sizes are again expected to be higher after two years, well above the minimum detectable probability of success according to our power calculations.

7.3. Offer Letters

***CCT* Offer Letter**

The Zomba Cash Transfer Program (ZCTP) with funding from the World Bank, would like to offer you, ___[NAME]___, a cash transfer to help you and your family with the burdens of school attendance for the 2008 school year. By accepting this offer, in return for going to school you will be given ___[AMOUNT]___ kwacha per month. If you attend secondary school, your fees for the 2008 school year will be paid directly to your school at the beginning of each term.

The payments to you and your family will be made on a monthly basis beginning in February, and will continue for 10 months through November 2008.

You are receiving this money in order to help you return to school or stay in school. In order to receive this money you **MUST** attend school at least **80% of the days for which your school is in session**.

***UCT* Offer Letter**

The Zomba Cash Transfer Program (ZCTP), with funding from the World Bank, would like to offer you, ___[NAME]___, a cash transfer to help you and your family. By accepting this offer you will be given ___[AMOUNT]___ kwacha per month.

The payments to you and your family will be made on a monthly basis beginning in February, and will continue for 10 months through November 2008.

These monthly transfer amounts specified above are given to you as a result of a lottery. You are not required to do anything more to receive this money. You will receive this money for 10 months between February and November, 2008.

7.4. Pre-Analysis Plan

PRE-ANALYSIS PLAN FOR ROUND 4 SCHOOLING, INCOME, AND HEALTH RISK IN MALAWI (SIHR) DATA
(SUB-SET OF VARIABLES USED IN THIS ANALYSIS)

Principal Investigators:

Sarah Baird, University of Otago
Ephraim Chirwa, Chancellor College
Craig McIntosh, UCSD
Berk Özler, World Bank and University of Otago

Analysis plan:

The core analysis will compare the impact of the Conditional Cash Transfer (CCT) and Unconditional Cash Transfer (UCT) treatment to control EAs for the baseline schoolgirl stratum, and will compare the CCT treatment to the control for the baseline dropout stratum. Most of the analysis will consist of Round 4 cross-sectional regression (using OLS unless not appropriate), although where possible we will also pursue panel difference-in-differences analysis for variables that have been consistently collected in multiple rounds. For consistency, the analysis will include the full set of controls used in the paper Baird, McIntosh and Özler (2011). These controls include baseline values of the following: a household asset index, highest grade attended, a dummy variable for having started sexual activity, dummy variables for age, and strata dummies. Standard errors will be clustered at the EA level, and results will be weighted to make them representative of the target population in the study EAs.

Only if a significant impact is found in the core analysis will further heterogeneity of impact be explored. Heterogeneity will be explored along both experimental dimensions including the amount of the transfer and the split between the parent and the girl, as well as based on the age at which you got the program and differences between rural and urban.

Note on construction of indexes:

To construct indexes for classes of variables, we will adhere to the following rubric:

a) For each sub-question in a family of variables, first align answers so that higher numbers always have a consistent meaning (good or bad).
b) Calculate the mean and SD of the responses to each sub-question in the sample *in the control group – separately for baseline schoolgirls and baseline dropouts*.
c) Create normalized variables that have the mean subtracted off and are divided by the SD.
d) Calculate the raw mean of the normalized variables for all sub-questions within a family of variables. This mean is the 'index' for those variables. *This summary index can further be normalized if desired.*

For the core analysis we will not pursue the analysis of sub-variables within an index unless the index as a whole is significant.

<u>Construction of Core Indexes: Primary outcomes are indicated in bold text.</u>

1. <u>Core Respondent-level outcomes:</u> Can be analysis with simple cross-sections comparing UCT, CCT, and control. No extensive margin issue with any of these variables.
    a. Schooling and Marriage (replication of QJE results with age-appropriate dependent variables):
        i. **Highest grade completed (S7, Q7)**
        ii. Highest educational qualification achieved (S7, Q9)
        iii. **Achievement, replacing the test scores with the 'competencies' (see below for construction). We will show the components, as well as the index for the quality index; and show only an index for the quantity.**
        iv. **Ever married (Part II CS, Q2e), ever pregnant (S18, Q1, Q2), number of live births (S18, Q17).**
        v. Hazard model of age of first marriage (S14, Q1 (and Round 3 data for those already married at Round 3)) and age at first birth (construct using age of respondent and DOB), with 'uncompleted spells' for those never married or never first birth.
        vi. Sexual behavior: ABC, # partners ever, as in previous papers. Ever had sex (S12, Q2, Q3, Q4) age at first sex (S12 Q4), total number of partners ever (S12, Q5), sexually active in past 12 months (S12, Q7), condom use last sex with most recent partner (S12, Q23)

    b. Health:
        i. Desired fertility S16 Q4 or Q10.
        ii. Mental health, calculated as in Baird, de Hoop, Ozler (2013) (S9, Q9-20): binary
        iii. Number of meals eaten with meat, eggs, fish in past 7 days (S9, Q 6-8)

Construction of Competencies Index:

Moderator variable for fertilizer application: S11bQ22

Fertilizer (Q23-26):

Quantity index: time taken to complete (Q23), categorize it as 1 'below median' (in seconds); 2 'above median'; and 3 'did not complete/did not complete in time'. Median is calculated among those who completed under the allocated time.

Quality index: Each Q (24-26) coded as 1 if Yes 0 if No and then added up to create an index between 0-3 of the quality of the application of fertilizer.

Normalize each by subtracting the control mean and dividing by the control SD.

Making change (Q27-28):

Same as above: (quantity index, Q27) and quality index (Q28).

Use the same procedure for Q29-30, Q31-32. Then, add the quality indices (Q28, 30, and 32). Add quantity indices (Q27, 29, and 31).

Normalize each by subtracting the control mean and dividing by the control SD.

Sending a text message (Q35-37) – moderator variables to be used for adjustment (Q33-34):

Same as above, then normalize each index.

Use the calculator on mobile phone (Q38-39):

Same as above, then normalize each index.

Calculate profits from trade (Q-40-42):

Same as above, then normalize each index.

Finally, average the normalized quantity indices and the quality indices separately to produce two final competency indices.

7.5. Survey Instruments

All survey instruments for all four rounds of data collection can be obtained from the authors upon request. We also replicate the modules for the competencies below.

SECTION 11b: EDUCATION - COMPETENCIES

Now I am going to ask you a series of questions that test your abilities in a variety of areas, including following instructions, mathematic skills, and reading and listening comprehension. This round there is no separate testing team, so we will complete all testing now. We will do a few story problems, where I explain various scenarios to you and ask for your answers.

22.	First, have you ever received agricultural extension training on how to apply fertilizer?	Yes.......... 1 No 2	[_____]

I am going to describe to you the process for optimum application of fertilizer to maize. Please listen very carefully because I am going to tell you this only once and you will need to remember how to complete the task. Please do not ask questions; I will only give these instructions once. Then I will time how long it takes you to complete it successfully. Are you ready to start?

READ THIS PART EXACTLY AS IT IS HERE, PAUSING FOR A MOMENT AFTER EACH SENTENCE:

> Take two tablespoons 23:21 and one tablespoon Can and mix them in the plate provided.
> Then, using the spoons and ruler provided, put one tablespoon of the resulting mixture
> 5 centimeters on either side of the root of the middle plant within the row of maize.

START THE TIMER - DO NOT REPEAT THE INSTRUCTIONS AGAIN.

23.	Time taken to complete (MM:SS)	[___|___] : [___|___] **MAX OF 5 MIN**	
24.	Is the mixture correct?	1= Yes 2= No	[_____]
25.	Was the correct spoon used?	1= Yes 2= No	[_____]
26.	Was the mixture placed within 4-6 cm of the "plant" and on both sides?	1= Yes 2= No	[_____]

GIVE THE RESPONDENT A PENCIL & THE ANSWER SHEET FACE DOWN TO BE USED AS SCRATCH PAPER.

You can use this page as scratch paper - please do not flip it over. I want you to imagine that you are an entrepreneur running a business. For the purpose of this exercise, I am going to provide you with change that you should consider to be the till in your business. **[GIVE CR THE CHANGE.]** I will now present you with different scenarios, and wish you to give me the correct change under each scenario. If needed, you may ask me to repeat information, but please work as quickly as you can.

SCENARIO A:

START THE TIMER THEN BEGIN TO READ THE QUESTION ALOUD.

READ THIS PART EXACTLY AS IT IS HERE, PAUSING FOR A MOMENT AFTER EACH SENTENCE:

> Assume I am a customer coming to your stall. I buy 3 tomatoes, which cost K20 each.
> I buy 2 bunches of onions, which cost K40 each. I also buy a box of matches that costs K18.
> I give you a K200 note. Please make my change.

Remind respondent that you can repeat the scenario. Keep timer running.

27.	Time taken to complete (MM:SS)	[___|___] : [___|___] **MAX OF 4 MIN**
28.	Amount of change given	MK [___|___|___]

SECTION 11b (cont'd): EDUCATION - COMPETENCIES

SCENARIO B

Now I would like you to do a similar scenario. This time, instead of me reading aloud to you, you will read the story yourself. Flip the paper over - this is Story B [POINT TO IT IF NECESSARY]. Again, please work as quickly as you can.

START TIMER AS THE RESPONDENT BEGINS TO READ THE QUESTION.

Do not help her read the story. Do not answer any questions related to the story.

29.	Time taken to complete (MM:SS)	[__|__] : [__|__] **MAX OF 4 MIN**
30.	Amount of change given	MK [__|__|__|__]

SCENARIO C

Now please continue to the next scenario. [POINT TO SCENARIO C IF NECESSARY] again, please read the story to youself and work as quickly as you can.

START TIMER AS THE RESPONDENT BEGINS TO READ THE QUESTION.

31.	Time taken to complete (MM:SS)	[__|__] : [__|__] **MAX OF 4 MIN**
32.	Amount of change given	MK [__|__|__|__]

Now I want to understand a bit about your familiarity with mobile phones.

33.	Have you ever used a mobile phone to send a text?	1= Yes 2= No >> 35	[____]
34.	Have you ever used THIS MODEL of mobile phone?	1= Yes 2= No	[____]

Please use this phone to send a text message with the word 'Hello' to the following number [SHOW WHERE THE INSTRUCTIONS AND PHONE NUMBER ARE ON THE PAGE]

START TIMER AS SOON AS YOU GIVE INSTRUCTIONS ONCE.

Remind respondent that you can repeat the phone number or the instructions - keep timer running.

35.	Time taken to complete (MM:SS)		[__|__] : [__|__] **MAX OF 4 MIN**	
36.	Was the message successfully sent?	1= Yes	2= No	[____]
37.	Was 'Hello' spelled correctly? (not case sensitive)	1= Yes	2= No	[____]

Now we will do the next item on your paper - use the calculator on this phone to calculate what is 873 x 17

START TIMER AS SOON AS YOU GIVE INSTRUCTIONS ONCE.

Remind respondent that you can repeat the calculation or the instructions - keep timer running.

38.	Time taken to complete (MM:SS)	[__|__] : [__|__] **MAX OF 4 MIN**
39.	What was the number given?	[__|__|__|__|__|__]

Now I am going to explain a farming scenario and ask you a few questions about it. You can follow along on the page in front of you, and may continue to use it as scrap paper as needed.

START TIMER THEN BEGIN TO READ THE QUESTION ALOUD.

Remind respondent that you can repeat the calculation or the instructions - keep timer running

In your village, a 50 kg bag of maize is selling for MKw 3,000. In a nearby community the same bag of maize is instead selling for 3,100. The transport to and from the nearby community will cost you Mkw 1,350 and you have 16 bags of maize to sell.

40.	Which one makes more profit - selling in your village, or in the nearby community? KEEP TIMER RUNNING	Selling in home village............1 Selling in nearby community....2 Won't answer ....3 >>42
41.	How much more profit does it make?	[____________]
42.	Time taken to complete (MM:SS)	[__|__] : [__|__] **MAX OF 4 MIN**

7.6. Analytical Tables and Results

Table 2: Attrition

	Baseline Schoolgirl	Baseline Dropout
	HH Panel	HH Panel
	(1)	(3)
=1 if Conditional	0.054***	0.011
	(0.018)	(0.035)
=1 if Unconditional	0.046*	
	(0.025)	
p-value UCT vs. CCT	0.766	N/A
p-value Treatment	0.008	N/A
Mean in Control Group	0.864	0.806
Number of observations	2,284	889

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 3: Baseline Means and Balance

	Baseline Schoolgirl Mean (s.d.)			p-value (CCT-UCT)	Baseline Dropout Mean (s.d.)	
	Control group	Conditional group	Unconditional Group		Control group	Conditional group
Urban Household	0.348	0.475	0.427	0.783	0.181	0.126
	(0.477)	(0.500)	(0.496)		(0.385)	(0.333)
Mother Alive	0.841	0.798	0.834	0.304	0.786	0.754
	(0.366)	(0.402)	(0.373)		(0.410)	(0.431)
Father Alive	0.71	0.716	0.767	0.238	0.659	0.651
	(0.454)	(0.451)	(0.424)		(0.475)	(0.477)
Household Size	6.38	6.349	6.664	0.168	6.118	6.138
	(2.265)	(2.145)	(2.070)		(2.403)	(2.623)
Asset Index	0.637	1.063	1.342*	0.563	-0.806	-0.722
	(2.579)	(2.709)	(2.433)		(2.246)	(2.487)
Age	15.219	14.911*	15.433	0.004	17.622	17.188
	(1.897)	(1.826)	(1.918)		(2.385)	(2.493)
Highest Grade Attended	7.498	7.242	7.906**	0.005	6.142	5.955
	(1.646)	(1.599)	(1.580)		(2.857)	(2.877)
Never Had Sex	0.803	0.806	0.786	0.604	0.305	0.293
	(0.398)	(0.395)	(0.411)		(0.461)	(0.456)
Ever Pregnant	0.021	0.030	0.030	0.981	0.447	0.417
	(0.144)	(0.170)	(0.170)		(0.498)	(0.494)

Notes: Mean differences statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence. Stars on the coefficients in columns (2) and (3) indicate significantly different than the control group for baseline schoolgirls. Stars on the coefficients in column (6) indicate significantly different than the control group for baseline dropouts. Means are weighted to make them representative of the target population in the study EAs.

Table 4A: Education Outcomes (Baseline Schoolgirls)

	Highest Grade Completed			Highest Education Qualification			=1 if Passed Primary School (PSLC)			=1 if Passed Junior Secondary School (JCE)		
	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4
=1 if Conditional Schoolgirl	0.074	0.126*	0.132	0.011	0.073*	0.019	0.030	0.013	-0.013	-0.012	0.055**	0.033
	(0.090)	(0.069)	(0.081)	(0.055)	(0.037)	(0.047)	(0.039)	(0.024)	(0.019)	(0.022)	(0.028)	(0.028)
=1 if Unconditional Schoolgirl	0.123	0.103	0.107	0.039	0.038	-0.032	0.044	0.030	0.016	0.003	0.016	0.014
	(0.110)	(0.121)	(0.131)	(0.047)	(0.057)	(0.054)	(0.038)	(0.026)	(0.016)	(0.022)	(0.045)	(0.036)
p-value UCT vs. CCT	0.675	0.854	0.850	0.671	0.581	0.393	0.775	0.600	0.185	0.560	0.439	0.643
p-value Treatment	0.482	0.174	0.246	0.707	0.143	0.693	0.415	0.488	0.389	0.814	0.148	0.477
Mean in Control Group	8.581	9.677	10.416	0.643	1.168	1.603	0.494	0.776	0.880	0.143	0.337	0.538
Sample Size	1,942	2,019	2,019	1,944	2,019	2,017	1,944	2,019	2,017	1,944	2,019	2,017

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Note that in Round 2 and Round 3 highest grade completed is actually highest grade attended. Highest education qualification takes on a value from 0-4 where 0 if no education qualification and 4 is a post-secondary qualification. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 4B: Education Outcomes (Baseline Dropouts)

	Highest Grade Completed			Highest Education Qualification			=1 if Passed Primary School (PSLC)			=1 if Passed Junior Secondary School (JCE)		
	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4
=1 if Conditional Schoolgirl	0.570***	0.558***	0.615***	0.054	0.108***	0.123***	0.036	0.058**	0.079***	0.014	0.049**	0.029
	(0.075)	(0.102)	(0.125)	(0.038)	(0.040)	(0.040)	(0.025)	(0.025)	(0.027)	(0.019)	(0.021)	(0.023)
Mean in Control Group	6.365	6.967	7.038	0.424	0.504	0.545	0.327	0.351	0.373	0.085	0.123	0.140
Sample Size	679	718	718	678	718	718	678	718	718	678	718	718

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Note that in Round 2 and Round 3 highest grade completed is actually highest grade attended. Highest education qualification takes on a value from 0-4 where 0 if no education qualification and 4 is a post-secondary qualification. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 5A: Learning and Skills (Baseline Schoolgirls)

	Test Scores (Round 3 Only)			Competencies (Round 4 Only)						
	Cognitive	Math	English	Total Competency	Fertilizer	Change Given	Text Message	Calculator	Profit	Total Time
=1 if Conditional Schoolgirl	0.183***	0.094	0.149***	0.071	0.025	0.051	0.073	0.058	0.004	-0.118
	(0.050)	(0.062)	(0.057)	(0.058)	(0.072)	(0.073)	(0.071)	(0.055)	(0.076)	(0.086)
=1 if Unconditional Schoolgirl	0.099	0.013	-0.066	0.089	0.087	-0.013	0.147*	0.089	-0.045	-0.123
	(0.130)	(0.100)	(0.091)	(0.067)	(0.093)	(0.058)	(0.080)	(0.066)	(0.091)	(0.087)
p-value UCT vs. CCT	0.531	0.439	0.038	0.788	0.500	0.404	0.429	0.674	0.563	0.962
p-value Treatment	0.002	0.305	0.022	0.317	0.644	0.695	0.154	0.322	0.833	0.235
Mean in Control Group	0.007	0.008	-0.014	0.004	0.001	0.000	0.004	0.005	0.000	-0.002
Sample Size	1,995	1,995	1,995	2,019	2,019	2,017	2,018	2,018	2,019	2,019

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. All outcome variables are standardized. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 5B: Learning and Skills (Baseline Dropouts)

	Test Scores (Round 3 Only)			Competencies (Round 4 Only)						
	Cognitive	Math	English	Total Competency	Fertilizer	Change Given	Text Message	Calculator	Profit	Total Time
=1 if Conditional Schoolgirl	0.155**	0.127*	0.074	0.066	-0.052	-0.011	0.107	0.076	0.089	-0.011
	(0.068)	(0.066)	(0.069)	(0.058)	(0.071)	(0.063)	(0.072)	(0.072)	(0.077)	(0.093)
Mean in Control Group	-0.026	-0.004	-0.009	0.009	0.009	0.008	-0.003	0.005	0.010	0.018
Sample Size	703	703	703	716	716	715	715	715	716	716

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. All outcome variables are standardized. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 6A: Marriage and Fertility (Baseline Schoolgirls)

Panel A: Marriage

	Ever Married			**Age of First Marriage**		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	0.001	-0.010	-0.037	-0.401	-0.182	-0.005
	(0.012)	(0.024)	(0.028)	(0.282)	(0.151)	(0.149)
=1 if Unconditional Schoolgirl	-0.030**	-0.083***	-0.012	-0.404	-0.011	0.502**
	(0.012)	(0.024)	(0.048)	(0.330)	(0.288)	(0.205)
p-value UCT vs. CCT	0.036	0.018	0.633	0.994	0.578	0.032
p-value Treatment	0.036	0.004	0.417	0.265	0.479	0.048
Mean in Control Group	0.046	0.180	0.399	17.622	17.988	18.631
Sample Size	1944	2018	2019	101	390	805

Panel B: Fertility (Baseline Schoolgirls)

	Ever Pregnant			**Total Live Births**			**Age at First Live**		**Desired Fertility**		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	0.008	0.027	-0.026	0.021	0.003	0.018	-0.049	-0.139	-0.100	-0.165	-0.079
	(0.015)	(0.027)	(0.034)	(0.014)	(0.022)	(0.037)	(0.175)	(0.137)	(0.105)	(0.102)	(0.066)
=1 if Unconditional Schoolgirl	-0.010	-0.063**	-0.004	0.013	-0.055*	-0.024	-0.193	-0.004	0.095	0.045	-0.017
	(0.017)	(0.028)	(0.043)	(0.017)	(0.030)	(0.047)	(0.229)	(0.169)	(0.123)	(0.105)	(0.057)
p-value UCT vs. CCT	0.358	0.009	0.643	0.669	0.075	0.447	0.584	0.470	0.201	0.121	0.433
p-value Treatment	0.655	0.025	0.744	0.257	0.151	0.745	0.698	0.580	0.429	0.217	0.497
Mean in Control Group	0.090	0.247	0.500	0.055	0.199	0.509	17.890	18.709	2.811	2.901	2.974
Sample Size	1943	2019	2019	1943	2019	2019	436	983	1944	2012	2018

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 6B: Marriage and Fertility (Baseline Dropouts)

Panel A: Marriage

	Ever Married			**Age of First Marriage**		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	-0.142***	-0.157***	-0.103***	0.024	0.255	0.408**
	(0.029)	(0.037)	(0.033)	(0.273)	(0.174)	(0.162)
Mean in Control Group	0.292	0.575	0.800	19.000	19.200	19.684
Sample Size	679	718	718	142	334	484

Panel B: Fertility (Baseline Dropouts)

	Ever Pregnant			**Total Live Births**			**Age First Live**		**Desired Fertility**		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	-0.058*	-0.081***	-0.038*	-0.010	-0.095**	-0.152***	0.237	0.313*	-0.110	-0.030	-0.160*
	(0.031)	(0.027)	(0.021)	(0.034)	(0.044)	(0.055)	(0.153)	(0.166)	(0.106)	(0.083)	(0.093)
Mean in Control Group	0.614	0.784	0.921	0.526	0.819	1.386	18.187	18.478	2.953	2.975	3.222
Sample Size	679	718	718	679	718	718	463	611	678	716	718

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 7A: Health (Baseline Schoolgirls)

	=1 if Suffers from Psychological Distress			Number of Meals Eaten		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	-0.071**	-0.037	-0.030	0.395**	0.596***	0.070
	(0.032)	(0.047)	(0.033)	(0.191)	(0.174)	(0.142)
=1 if Unconditional Schoolgirl	-0.141***	-0.026	-0.009	0.448**	0.338**	-0.088
	(0.036)	(0.054)	(0.048)	(0.199)	(0.153)	(0.255)
p-value UCT vs. CCT	0.077	0.860	0.668	0.835	0.215	0.578
p-value Treatment	0.000	0.677	0.662	0.020	0.001	0.825
Mean in Control Group	0.373	0.313	0.368	3.971	4.052	4.146
Sample Size	1,940	2,013	2,015	1,944	2,018	2,017

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 7B: Health (Baseline Dropouts)

	=1 if Suffers from Psychological Distress			Number of Meals Eaten		
	Round 2	**Round 3**	**Round 4**	**Round 2**	**Round 3**	**Round 4**
=1 if Conditional Schoolgirl	0.002	0.010	0.019	0.358*	0.224	0.275
	(0.040)	(0.036)	(0.042)	(0.209)	(0.192)	(0.180)
Mean in Control Group	0.465	0.314	0.436	3.646	3.989	3.715
Sample Size	679	715	717	679	718	718

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 8A: Sexual Behaviour (Extensive Margin) (Baseline Schoolgirls)

	Ever Had Sex			# Sexual Partners			Sexually Active Past 12 Months		
	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4
=1 if Conditional Schoolgirl	0.007	0.005	0.004	-0.021	0.005	-0.013	-0.009	0.001	-0.034
	(0.025)	(0.029)	(0.039)	(0.040)	(0.048)	(0.065)	(0.023)	(0.029)	(0.037)
=1 if Unconditional Schoolgirl	-0.004	0.020	0.045	-0.037	-0.007	0.102	-0.018	-0.036	0.039
	(0.032)	(0.030)	(0.039)	(0.048)	(0.036)	(0.068)	(0.030)	(0.032)	(0.045)
p-value UCT vs. CCT	0.760	0.692	0.418	0.772	0.815	0.128	0.775	0.327	0.162
p-value Treatment	0.947	0.810	0.519	0.706	0.969	0.239	0.803	0.514	0.367
Mean in Control Group	0.235	0.398	0.684	0.329	0.559	1.042	0.174	0.308	0.561
Sample Size	1,942	2,016	2,018	1,941	2,016	2,017	1,942	2,015	2,018

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 8B: Sexual Behaviour (extensive margin) (Baseline Dropouts)

	Ever Had Sex			# Sexual Partners			Sexually Active Past 12 Months		
	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4	Round 2	Round 3	Round 4
=1 if Conditional Schoolgirl	-0.047**	-0.067***	0.001	0.008	-0.118	-0.016	-0.122***	-0.094**	-0.040
	(0.024)	(0.025)	(0.012)	(0.159)	(0.153)	(0.097)	(0.035)	(0.037)	(0.029)
Mean in Control Group	0.787	0.910	0.970	1.409	1.734	2.066	0.506	0.674	0.825
Sample Size	679	718	718	679	718	718	679	718	718

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 9A: Sexual Behaviour (intensive margin) (Baseline Schoolgirls)

	Age at First Sex			Condom Use	
	Round 2	Round 3	Round 4	Round 3	Round 4
=1 if Conditional Schoolgirl	0.189	0.136	0.187	-0.006	0.010
	(0.138)	(0.130)	(0.146)	(0.055)	(0.042)
=1 if Unconditional Schoolgirl	-0.195	-0.039	-0.198	0.102	0.056
	(0.166)	(0.189)	(0.131)	(0.086)	(0.049)
p-value UCT vs. CCT	0.050	0.404	0.037	0.268	0.454
p-value Treatment	0.137	0.536	0.108	0.483	0.518
Mean in Control Group	15.748	16.393	17.186	0.247	0.270
Sample Size	516	893	1,469	672	1,161

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. 'Age at First Sex' is defined for those that had ever had sex. 'Older Partner' is defined as having a partner who is 5 years older or more in the past 12 months. 'Condom Use' is defined as using a condom at last sex with most recent sexual partner. It is missing for those who were not sexually active in the past 12 months. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table 9B: Sexual Behaviour (intensive margin)(Baseline Dropouts)

	Age at First Sex			Condom Use	
	Round 2	Round 3	Round 4	Round 3	Round 4
=1 if Conditional Schoolgirl	-0.013	-0.061	0.102	0.046	0.031
	(0.137)	(0.144)	(0.135)	(0.037)	(0.031)
Mean in Control Group	16.234	16.578	16.777	0.159	0.161
Sample Size	513	625	697	446	577

Notes: Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make them representative of the target population in the study EAs. Baseline values of the following variables are included as controls in the regression analyses: age dummies, strata dummies, household asset index, highest grade attended, and an indicator for never had sex. We restrict the sample to respondents who were surveyed in both Round 3 and Round 4. 'Age at First Sex' is defined for those that had ever had sex. 'Older Partner' is defined as having a partner who is 5 years older or more in the past 12 months. 'Condom Use' is defined as using a condom at last sex with most recent sexual partner. It is missing for those who were not sexually active in the past 12 months. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.