# CASH OR CONDITION?
# EVIDENCE FROM A CASH TRANSFER EXPERIMENT[1]

SARAH BAIRD

CRAIG MCINTOSH

BERK ÖZLER

February 23, 2011

**Abstract**

This paper assesses the role of conditionality in cash transfer programs using a unique experiment targeted at adolescent girls in Malawi. The program featured two distinct interventions: unconditional transfers (UCT arm) and transfers conditional on school attendance (CCT arm). While there was a modest decline in the dropout rate in the UCT arm in comparison to the control group, it was only 43% as large as the impact in the CCT arm at the end of the two-year program. The CCT arm also outperformed the UCT arm in tests of English reading comprehension. However, teenage pregnancy and marriage rates were substantially lower in the UCT than the CCT arm, entirely due to the impact of UCTs on these outcomes among girls who dropped out of school.

Keywords: Conditional Cash Transfers, Education, Adolescent Girls, Fertility

JEL Codes: C93, I21, I38, J12

---

[1] We are grateful to four anonymous referees for helpful feedback on earlier drafts of this paper, as well as seminar participants at CEGA, George Washington University, IFPRI, NEUDC, Paris School of Economics, Toulouse School of Economics, UC Berkeley, UC San Diego, University of Namur, University of Pennsylvania, and the World Bank for useful discussions. We particularly appreciate the numerous discussions we had with Francisco Ferreira on this topic. We thank everyone who provided this project with great field work and research assistance and are too numerous to list individually. We gratefully acknowledge funding from the Global Development Network, the Bill and Melinda Gates Foundation, NBER Africa Project, World Bank Research Support Budget Grant, as well as several trust funds at the World Bank: Knowledge for Change Trust Fund (TF090932), World Development Report 2007 Small Grants Fund (TF055926), Spanish Impact Evaluation Fund (TF092384), and Gender Action Plan Trust Fund (TF092029). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development or the World Bank. Please send correspondence to: sbaird@gwu.edu, ctmcintosh@ucsd.edu, or bozler@worldbank.org.

## 1. INTRODUCTION

Conditional Cash Transfers (CCTs) are "… targeted to the poor and made conditional on certain behaviors of recipient households" (World Bank 2009). A large and empirically well-identified body of evidence has demonstrated the ability of CCTs to improve schooling outcomes in the developing world (Schultz 2004; de Janvry et al. 2006; Filmer and Schady 2010; among many others).[2] Due in large part to the high-quality evaluation of Mexico's PROGRESA, CCT programs have become common in Latin America and began to spread to other parts of the world, with CCT programs now in more than 29 developing countries (World Bank 2009).[3]

There are also rigorous evaluations of Unconditional Cash Transfers (UCTs), which cover a wide range of programs: non-contributory pension schemes, disability benefits, child allowance, and income support. Whether examining a cash transfer program in Ecuador (Bono de Desarollo Humano), the old age pension program in South Africa, or the child support grants also in South Africa, studies find that the UCTs reduce child labor, increase schooling, and improve child health and nutrition (Edmonds and Schady 2009; Edmonds 2006; Case, Hosegood, and Lund 2005; Duflo 2003).[4] Hence, UCTs also change the behaviors on which CCTs are typically conditioned.

The debate over the relative merits of these two approaches has intensified as CCTs have become more widely implemented. Proponents of CCT programs argue that market failures may often lead to underinvestment in education or health, which are addressed by the conditions imposed on recipient households. Another advantage of CCT programs is that the conditions make cash transfers politically palatable to the middle and upper class voters who are not direct

---

[2] Many CCT programs require households to comply with only schooling conditions, while some also require compliance with health conditions, such as regular visits to health clinics for children.

[3] CCTs are also implemented in developed countries. For example, a three-year pilot CCT program in New York City ended in early 2010. For more on Opportunity NYC, see: http://www.nyc.gov/html/ceo/html/programs/opportunity_nyc.shtml.

[4] For a recent review of cash transfer programs, see Adato and Bassett (2009), which gives more examples of unconditional cash transfer programs in Sub-Saharan Africa improving education, health, and nutrition among children.

beneficiaries of such programs.[5] To critics of conditionality, the 'theoretical default' position should be to favor UCTs, particularly because the marginal contribution of the conditions to cash transfer programs remains largely unknown.[6] Furthermore, the implementation of CCT programs may strain administrative capacity as these programs expand to poorer countries outside of Latin America.

The existing knowledge base concerning the marginal impact of attaching conditions to cash transfer programs remains very limited – especially in sub-Saharan Africa where such evaluations are relatively rare.[7] One strain of relevant literature relies on accidental glitches in program implementation. Based on the fact that some households in Mexico and Ecuador did not think that the cash transfer program in their respective country was conditional on school attendance, de Brauw and Hoddinott (2010) and Schady and Araujo (2008) both find that school enrollment was significantly lower among those who thought that the cash transfers were *unconditional*. There is also a literature that takes a structural approach, where a model of household behavior is calibrated using real data, and then the impacts of various policy experiments are simulated. Micro-simulating Brazil's Bolsa Escola program, Bourguignon, Ferreira, and Leite (2003) find that UCTs would have no impact on school enrollment. Todd and Wolpin (2006), examining PROGRESA in Mexico, report that the increase in schooling with unconditional transfers would be only 20% as large as the conditional transfers while the cost per family would be an order of magnitude larger.[8] Overall, the

[5] For an excellent discussion of the economic rationale for conditional cash transfers, see Chapter 2 in World Bank (2009).

[6] For a discussion of "The Conditionality Dilemma", see Chapter 8 in Hanlon, Barrientos, and Hulme (2010).

[7] A few experiments to improve the design of CCT programs have recently been conducted, most notably in Colombia (Barrera-Osorio et al., forthcoming).

[8] There is also some evidence that the condition that pre-school children receive regular check-ups at health clinics (enforced by a social marketing campaign, but not monitoring the condition) had a significant impact on child cognitive outcomes, physical health, and fine motor control. Two studies in Latin America – Paxson and Schady (2007) and Macours, Schady, and Vakis (2008) – show behavioral changes in the spending patterns of parents and households that they argue to be inconsistent with changes in *just* the household income. These studies, however, cannot isolate the impact of the social marketing campaign from that of the transfers being made to women.

existing non-experimental evidence suggests that conditionality plays an important role in the overall impact of CCTs.[9]

The ideal experiment to identify the marginal contribution of conditionality in a cash transfer program – i.e. a randomized controlled trial with one treatment arm receiving *conditional* cash transfers, another receiving *unconditional* transfers, and a control group receiving *no* transfers – has not previously been conducted anywhere.[10] This paper describes the impacts of such an experiment in Malawi that provided cash transfers to households with school-age girls. In the experiment, 176 enumeration areas (EAs) were randomly assigned treatment or control status.[11] A sub-group of the 88 treatment EAs was then randomly assigned to receive offers for monthly cash transfers *conditional* on attending school regularly (CCT arm) while another group of EAs received offers for *unconditional* cash transfers (UCT arm). Offers included separate transfers to the girls and their parents/guardians. The transfer amounts offered to the parents were randomized at the village level, and those offered to the girls were randomized at the individual level.

In this paper, we exploit this experiment to not only examine the impact of each treatment arm on behaviors on which the CCT intervention was conditioned (school enrollment and attendance), but also on outcomes that are of central importance to the long-term prospects of school-age girls: human capital formation (measured by tests of English reading comprehension, mathematics, and cognitive skills), marriage, and childbearing. At the micro-level, improved test scores are associated with increased wages later in life (Blau and Kahn 2005), while delayed fertility is

---

[9] However, using an experiment that provided in-kind food transfers in one arm and equal-valued unrestricted cash transfers in another arm in Mexico, Cunha (2010) finds that households receiving the latter consumed equally nutritious foods as the former and that there were no differences in anthropometric or health outcomes of children between the two treatment arms. The study concludes that there is little evidence to justify the paternalistic motivation when it comes to this in-kind food transfer program.

[10] To our knowledge, there are two other studies that plan to examine the impact of the conditionality in the near future. "Impact Evaluation of a Randomized Conditional Cash Transfer Program in Rural Education in Morocco" has three treatment arms: unconditional, conditional with minimal monitoring, and conditional with heavy monitoring (using finger printing machines at schools). A similar pilot in Burkina Faso has comparative treatment arms for conditional and unconditional transfers.

[11] An EA consists of approximately 250 households spanning several villages.

associated with improved maternal and child health outcomes.[12] Increased age at first marriage can improve the quality of marriage matches and reduce the likelihood of divorce, increase women's decision-making power in the household, reduce their chances of experiencing domestic violence, and improve health care practices among pregnant women (Goldin and Katz 2002; Jensen and Thornton 2003; Field and Ambrus 2008). At the macro level, improved cognition may lead to more growth (Hanushek and Woessmann 2009), while lower fertility rates may also contribute to economic growth through increased female labor supply (Bloom et al. 2009) and by allowing greater investments in the health and education of children.

Starting with schooling outcomes, we find that while dropout rates declined in both treatment arms, the effect in the UCT arm is 43% as large as that in the CCT arm.[13] Evidence from school ledgers for students enrolled in school also suggests that the fraction of days attended in the CCT arm is higher than the UCT arm. Using independently administered tests of cognitive ability, mathematics, and English reading comprehension, we find that while achievement is significantly improved in all three tests in the CCT arm compared with the control group, no such gains are detectable in the UCT arm. The difference in program impacts between the two treatment arms is significant at the 90% confidence level for English reading comprehension. In summary, the CCT arm had a significant edge in terms of schooling outcomes over the UCT arm: a large gain in enrollment and a modest yet significant advantage in learning.

---

[12] Evidence on the causal effects of childbearing as a teen is inconclusive in both the biomedical and the economics literature. While some argue that gynecological immaturity increases the likelihood of preterm births and competition for nutrients between the mother and baby can cause low birth weight, the evidence is mixed (Fraser, Brockert, and Ward 1995; Akinbami, Schoendorf, and Kiely 2000; Smith and Pell 2001; Horgan and Kenny 2007). In economics, while Hotz, McElroy, and Sanders (2005) argue that "much of the 'concern' that has been registered regarding teenage childbearing is misplaced," the debate is ongoing (Ashcraft and Lang 2006; Fletcher and Wolfe 2008; Dahl, 2010).

[13] Our preferred measure of enrollment uses enrollment data for each term in 2008 and 2009, which are confirmed by the teachers during the school surveys. Self-reported measures of enrollment produce divergent results in terms of the relative effectiveness of the CCT and UCT arms. We discuss measurement issues in detail in Section 2 and present enrollment impacts using self-reported data as well as data collected from the schools in Section 3.

When we turn to examine marriage and pregnancy rates, however, unconditional transfers dominate. The likelihood of being 'ever pregnant' and 'ever married' were 27% and 44% lower in the UCT arm at the end of the two-year intervention, respectively, whereas program impacts on these two outcomes were small and statistically insignificant in the CCT arm. These substantial delays in marriage and fertility in the UCT arm are found entirely among adolescent girls who dropped out of school after the start of the two-year intervention; rates of marriage and fertility among girls still enrolled in school at follow-up were negligible regardless of treatment status. Hence, the success of the conditionality in promoting the formation of human capital among the compliers appears to be achieved at the cost of denying transfers to non-compliers who are shown to be particularly 'at risk' for early marriage and teenage pregnancy.

By exploiting an experiment featuring a CCT and a UCT arm and by broadening the impact assessment beyond schooling, our study exposes a trade-off that is inherent in CCT programs. The existing literature is focused primarily on assessing the desired behavior change in CCT programs, and may have overlooked the effects of denying benefits to those who fail to satisfy the conditions. Our findings show that UCTs can improve important outcomes among such households even though they might be much less effective than CCTs in achieving the desired behavior change. The trade-off between improved schooling outcomes and delayed marriage and childbearing among school-age girls illuminates the importance of carefully considering what exactly transfer programs are trying to achieve in the target population.

In the next section, we describe the study setting and sample selection; the research design and the intervention; as well as the multiple sources of data collection undertaken for this study. As the CCT and UCT interventions took place simultaneously in different communities within the same district, we include a discussion of the circumstances under which this experiment was conducted and provide evidence on the program beneficiaries' understanding of program rules. Issues concerning the measurement of various schooling outcomes are also discussed in this section. Section 3 describes the estimation strategy, presents the main program impacts on schooling, fertility, and marriage by treatment arm, as well as the heterogeneity of these impacts by age group and transfer size. Section 4 concludes.

## 2. BACKGROUND, STUDY DESIGN, AND DATA

### 2.1. Study setting

Malawi, the setting for this research project, is a small and poor country in southern Africa. 81% of its population of 15.3 million lived in rural areas in 2009, with most people relying on subsistence farming. The country is poor even by African standards: Malawi's 2008 GNI per capita figure of $760 (PPP, current international $) is less than 40% of the sub-Saharan African average of $1,973 (World Development Indicators Database, 2010). According to the same data source, net secondary school enrollment is very low at 24%.

### 2.2. Sampling

Zomba District in the Southern region was chosen as the site for this study. 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.[14] The target population was then divided into two main groups: those who were out of

[14] The target population of 13-22 year-old, never-married females was selected for a variety of reasons. As the study was designed with an eye to examine the possible effect of schooling cash transfer programs on the risk of HIV infection, we

school at baseline (*baseline dropouts*) and those who were in school at baseline (*baseline schoolgirls*). *Baseline schoolgirls*, who form 87% of the target population within our study EAs and among whom the 'conditionality' experiment was carried out, are the subject of this paper.[15] In each EA, a percentage of baseline schoolgirls were randomly selected for the study. These sampling percentages 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 2,907 schoolgirls in 176 EAs, or an average of 16.5 per EA.

### 2.3. Study Design and Intervention

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.[16] Excluding the 623 girls who lived in intervention EAs but did not receive an offer, we are left with a sample of 2,284 *baseline schoolgirls* in 161 EAs (1,495 in 88 control EAs, 506 in 46 CCT EAs, and the remaining 283 in 27 UCT EAs). See Figure I for an

---

focused on females because the HIV rate among boys and young men of schooling age is negligible. The age range was selected so that the study population was school-age 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.

[15] Many cash transfer programs are school-based, meaning that they do not cover those who have already dropped out of school (see, for example, the discussion of Cambodia's CESSP Scholarship Program in Filmer and Schady, 2009). Other programs, such as PROGRESA in Mexico, covered baseline dropouts, but studies usually exclude this group from the evaluation due to the 'one-time effect' at the onset of the program for this group (de Janvry and Sadoulet, 2006). While outcomes for *baseline dropouts* were also evaluated under the broader study, they are not the subject of this paper as the 'conditionality' experiment was not conducted among this group. As the sample size for this group is quite small (889 girls in 176 EAs at baseline, i.e. approximately 5 girls per EA), dividing the treatment group into a CCT and a UCT group would yield an experiment with low statistical power. Hence, in treatment EAs, this group received CCT offers only.

[16] 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.

illustration of the sample. No EA in the sample had a similar cash transfer program before or during the study.

**2.3.1. 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. 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.[17] An example of the CCT offer letters can be seen in Appendix A. 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 to $1, $2, $3, $4, or $5 per month.[18] 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.[19]

---

[17] Due to uncertainties regarding funding, the initial offers were only made for the 2008 school year. However, upon receipt of more funds for the intervention in April 2008, all the girls in the program were informed that the program would be extended to continue until the end of 2009.

[18] 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 has a Social Cash Transfer Scheme, which is now under consideration for scale-up at the national level that transferred $12/month plus bonuses for school-age children during its pilot phase (Miller and Tsoka 2007).

[19] 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.

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.[20] 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 they received and regardless of their 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.

#### 2.3.2. 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. An example of the UCT offer letters can also be seen in Appendix A. 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.[21] 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

---

[20] We were initially concerned that teachers may falsify attendance records for program beneficiaries – either out of benevolence for the student or perhaps to extract bribes. To ensure this did not happen, a series of spot checks were conducted about half way through the first year of the program in 2008. This meant that the program administrators went to a randomly selected sub-sample of schools attended by girls in the CCT arm and conducted roll calls for the whole class after attendance for that day had been taken. In all schools but one, the ledger perfectly matched the observed class attendance for that day. As these spot checks were expensive to conduct, they were discontinued after the study team was convinced that the school ledgers gave an accurate reflection of real attendance.

[21] Because the average school fees paid in the conditional treatment arm could not be calculated until the first term fees were paid, the adjustment in the unconditional treatment arm was made starting with the second of 10 monthly payments for the 2008 school year. The average school fees paid for secondary school girls in the conditional treatment group for Term 1 (3,000 Malawian Kwacha, or approximately $20) was multiplied by three (to calculate an estimate of the mean annual school fees), divided by nine (the number of remaining payments in 2008) and added to the transfers received by households with girls eligible to attend secondary school in the UCT arm. The NGO implementing the program was instructed to make no mention of school fees but only explain to these households that they were randomly selected to receive a 'bonus.'

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.[22] Even though the offer letters were differentiated carefully and treatment status for each individual was reinforced during the monthly cash transfer meetings by the implementing NGO, it is natural to question whether the beneficiaries in the UCT arm understood the program rules correctly. 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.

As summarized in Section 1 above and presented in detail in Section 3, we find statistically significant differences between programs impacts in the CCT and UCT for all the main outcome indicators examined in this paper: enrollment, test scores, marriage, and pregnancy. These differences offer *prima facie* evidence that the two interventions were perceived to be different than each other.[23]

In order to understand the perceptions of study participants more fully, we conducted qualitative interviews with a random sub-sample in the autumn of 2010 – approximately nine months after the two-year intervention ended in December 2009. Of the fifteen girls randomly selected from the UCT arm, only one of them reported a fear of losing payments if she was not attending school.[24] The interviews with those in the UCT arm lead to two clear conclusions. First, the rules of the program were well understood by the girls in the UCT arm; the interviews make

[22] The reader will note that there are no 'ideal conditions' under which to conduct this particular experiment. For example, we could have run the CCT and UCT experiments in separate districts with little inter-district communication, but then we would not be able to rule out the possibility of unobserved heterogeneity driving the results.
[23] In Appendix E, we provide further evidence on this issue by showing that the exogenous variation in the number of CCT beneficiaries at the monthly cash transfer meetings had no effect on outcomes in the UCT arm.
[24] Similarly, we can find no evidence supporting the notion that girls in the UCT arm thought their transfer payments were conditional on not getting married.

clear that UCT girls knew that nothing was required of them to participate in the program and they were given no rules or regulations tied to the receipt of the transfers other than showing up at the pre-determined cash transfer locations. Second, girls in the UCT arm were very much aware of the CCT intervention. Interviews suggest that the girls in the UCT arm not only knew about the CCT program, and but many actually had friends or acquaintances in the CCT arm. Through these contacts they knew that school attendance was strictly monitored in the CCT arm, and that non-compliers were penalized.[25]

The evidence from the in-depth interviews makes it clear that the UCT experiment did not happen in a vacuum. Instead, it 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.[26]

### 2.4. Data sources and outcomes

#### 2.4.1. Data sources

The data used in this paper were collected in three rounds. Baseline data, or Round 1, were collected between October 2007 and January 2008, before the offers to participate in the program took place. First follow-up data collection, or Round 2, was conducted approximately 12 months later – between October 2008 and February 2009. The second follow-up (Round 3) data collection was conducted between February and June 2010 – after the completion of the two-year intervention at the end of 2009 to examine the final impacts of the program. To assess program impacts on school enrollment, attendance, and achievement, as well as on fertility and marriage, we use multiple

[25]The following excerpt from an interview (with respondent ID 1461204) is a good example:
**Interviewer:** *Earlier you talked of conditional and unconditional. What did you say about the rules for conditional girls?*
**Respondent:** *They had to attend class all the time…not missing more than 3 days of classes in a [month] – like I already explained.*
**Interviewer:** *How did you say the program managers knew about the missed school days?*
**Respondent:** *They would go to the schools…For example, I have a friend, [name], who was learning at [school name]. They would go each month to the school to monitor her attendance, and if she was absent for more than three days she would not get her monthly money.*
[26] It is also possible that the existence of a CCT program could have reduced the motivation of the girls in the UCT arm.

data sources: household surveys (all rounds), school surveys (Rounds 2 and 3), school ledgers, independently developed achievement tests, and qualitative interviews (Round 3 only).

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.

During Round 2, we also 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 our study sample in 2008. This procedure was repeated in Round 3 for a randomly selected sub-sample of core respondents who reported being enrolled in school in 2009.[27] Using Round 2 (Round 3) household survey data, we collected the name of the school, the grade, and the teacher's name for the core respondent if she reported being enrolled in school at any point during the 2008 (2009) school year. These teachers were then located at the named schools and interviewed about each respondent's schooling status (term by term) during the past school year. Furthermore, during Round 3, school ledgers were sought to check the attendance of respondents for each school term in 2009 and the first term of 2010.

To measure program impacts on student achievement, *mathematics and English reading comprehension tests* were developed and administered to all study participants at their homes. The tests were developed by a team of experts at the Human Sciences Research Council according to the

[27] The reason why the school survey in Round 3 was conducted for a randomly selected sub-sample instead of every student who reported being enrolled in school in 2009 is that school ledgers were also sought to check the attendance of core respondents. As locating these ledgers, examining them, and recording attendance for each core respondent is time consuming and costly, the study team decided to reduce the sample size.

Malawian curricula for these subjects for Standards 5-8 and Forms 1-2.[28] 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).[29] The mathematics and English tests were piloted for a small, randomly selected sub-sample of the study participants in the control group before being finalized for administration during Round 3 data collection. These tests were administered by trained proctors at the residences of the study participants and were always administered after the household survey, but never on the same day. The order of the math and English tests were randomized at the individual level and the Raven's test was always administered last.

Finally, *structured in-depth interviews* were conducted with a small sample of study participants, their parents or guardians, community leaders, program managers, and schools. The sample was selected randomly using block stratification based on treatment status at baseline, as well as schooling and marital status at Round 3. The total number of baseline schoolgirls sampled was forty-eight, of whom forty-six were successfully interviewed.[30] The main aim of these structured interviews was to gauge the "understanding of the cash transfer intervention" by study participants. In addition, topics of discussion included schooling decisions, dating, fertility, and marriage, as well as empowerment and future aspirations. The interviews usually lasted sixty to ninety minutes and were conducted by trained enumerators, many of whom had previous experience in qualitative field work. The conversations in Chichewa were taped and transcribed in English immediately after the interviews.

[28] In Malawi, there are eight grades in primary school (Standard 1-8) and four in secondary school (Form 1-4).
[29] These three tests are available from the authors upon request.
[30] We have fourteen in-depth interviews in the control group, seventeen in the CCT arm, and fifteen in the UCT arm.

### 2.4.2. Outcomes

*Schooling*

We measure enrollment and attendance using three different data sources. The first indicator is constructed using self-reported data from the household survey on whether the core respondent was enrolled in school. These questions are asked for each of the seven school terms between Term 1, 2008 and Term 1, 2010. As self-reported data may overstate enrollment, we cross-validated these data by visiting the schools the study participants reported attending, and asking the same questions to the teachers of the core respondents. The enrollment indicators from the school survey are coded 'zero' if the core respondent reported not being enrolled in school for that term or if the teacher reported her as not enrolled and 'one' if her teacher(s) confirmed that she was attending school during the relevant term. Finally, as enrollment may be a poor proxy for actual school attendance, we utilize the attendance ledgers for the 2009 school year and the first term of 2010 collected during the school surveys in Round 3 to construct an indicator for the percentage of days the core respondent enrolled in school was recorded 'present' during days the school was in session.

We did not independently monitor the school attendance of study participants through random spot checks. While studies such as Miguel and Kremer (2004), Kremer, Miguel, and Thornton (2009), and Barrera-Osorio et al. (forthcoming) have measured attendance directly, we deliberately chose to forego this method of data collection to protect the validity of the UCT experiment. Despite having data on enrollment from the teachers and attendance from school ledgers, direct observation would clearly have produced superior evidence to the alternative measures of school participation used in this paper. However, as reported above in Section 2.3.2, girls in the UCT arm were fully aware that the attendance of CCT recipients were being regularly monitored, which led them to believe that program managers 'cared about the education' of the girls in the CCT arm. We were concerned that performing random spot checks of attendance for girls in the UCT arm could have given them the impression that they were also supposed to attend school regularly to receive their payments.[31] As a consequence, we chose to avoid direct monitoring of attendance in order to retain as sharp a differential test as possible of the relative merits of conditional and unconditional transfers.

As important as school attendance may be for adolescent girls, perhaps as important is learning achievement and cognitive skills.[32] To measure these, we conducted independently developed tests of mathematics, English reading comprehension, and cognitive ability. Total number of correct answers in each of these tests is standardized to have a mean equal to 'zero' and standard deviation equal to 'one' in the control group and program impacts are presented as changes in standard deviations (SD).

*Marriage and Fertility*

Teenage pregnancy in Malawi is common with the adolescent fertility rate at 133 per 1,000 women aged 15-19 (World Development Indicators 2010).[33] Many girls cite pregnancy as the main reason for dropping out of school and getting married at an early age. Each of the core respondents was asked the following questions in each round: "Have you ever been pregnant or are you currently pregnant?" and "What is your marital status?" We use the answers to these questions to calculate the prevalence of marriage and pregnancy in Rounds 2 and 3.

---

[31] This is nicely illustrated during an in-depth interview with one of the core respondents (respondent ID 1332203): After describing the CCT girls being followed to their schools to monitor their attendance, which she explained showed her that the program was interested in 'attracting girls to go to school,' she was asked by the interviewer in what way the program managers cared (about schooling). She answered by saying: "They cared for the conditional group only but on the other group they didn't care."

[32] Other schooling outcomes, such as repetition and reentry rates are also important and can lead to different inferences regarding program impacts on schooling attainment. See, for example, Behrman, Sengupta, and Todd (2005).

[33] For comparison, the same figure is 35 in the U.S.A. and 64 in Mexico (World Development Indicators 2010).

## 3. Estimation Strategy and Results

### 3.1. Sample Attrition and Balance

Figure I summarizes the study sample and attrition. We began with a sample of 2,284 respondents who were in school at baseline and formed the experimental sample for our study of conditionality. Of this sample, 2,186 were tracked successfully for the Round 3 household survey and 2,089 were successfully interviewed in all three rounds, a tracking rate of over 90%. Of the 983 subjects randomly sampled for the school survey in Round 3, enrollment data reported by teachers are available for 922 of them. We were less successful in locating attendance ledgers; of the 821 girls who were selected for the Round 3 school survey and reported being enrolled in school in 2009, legible ledgers for at least one term are only available for 384.

Table I examines attrition across the two treatment arms and the control group separately by each of our data sources: household surveys, achievement tests, and school surveys. The regression analysis indicates that there is no significant differential attrition between the two treatment arms in any of the six sub-samples used for analysis in this study. Study participants in both treatment arms, however, were equally more likely to take the achievement tests than the control group. Similarly, legible ledgers are more likely to be found for treatment girls. Thus, the analysis of the available samples should give us unbiased estimates of *differential* program impacts on schooling outcomes, marriage, and childbearing.[34]

In Table II, we test the balance of the experiment, using baseline data for the sample used in the analysis, i.e. those successfully interviewed during all three rounds. Panel A shows balance on household attributes, and Panel B on individual characteristics. Overall, the experiment appears well balanced between the treatment and control groups over a broad range of outcomes; Column (4) shows that the two treatment arms only differ in age and highest grade attended at baseline – two

[34] To assess the robustness of our findings with respect to attrition from the sample that took the achievement tests, we perform a bounding exercise (Lee, 2009), described in detail in footnote 40.

variables that are highly correlated with each other. While the share of female-headed households is balanced between the two treatment arms, it is significantly higher in the control group.[35]

### 3.2. Estimation Strategy

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 school enrollment, attendance, test scores, marriage, and pregnancy, we employ a simple reduced-form linear probability model of the following type:

$$Y_i = T_i^C \gamma^C + T_i^U \gamma^U + X_i \beta + \varepsilon_i, \tag{1}$$

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. The standard errors $\varepsilon_i$ 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.

[35] In Appendix Tables B.1 and B.2, we examine balance for the same set of characteristics for the school survey and ledger sub-samples analyzed in this paper. Both of these tables look similar to Table II, confirming balance in these sub-samples.

### 3.3. Results

### 3.3.1. Schooling

*Enrollment*

Table III describes enrollment rates by term, including a cumulative variable for the number of terms the girl was enrolled in school during the two-year intervention that takes on a value between 0 and 6. When we examine self-reported enrollment in Panel A, we see that dropout rates in the control group steadily increase over time with the sharpest change occurring between school years. Impact estimates suggest that dropout rates were significantly lower in both treatment arms, and that the UCT arm outperformed the CCT arm: the program impact on the number of terms the girls were enrolled in school during the two-year intervention is 0.41 terms in the UCT arm, compared to 0.23 terms in the CCT arm – a difference that is significant at the 95% confidence level.

Self-reported enrollment data can be subject to reporting bias. For example, comparing program impacts using self-reports vs. monitored data, Barrera-Osorio et al. (forthcoming) report that a significant positive bias in self-reported school enrollment compresses the difference between the treatment and control groups causing a downward bias in observed program impact. Baird and Özler (2010) confirm this finding for Malawi and show that differential misreporting can further bias program impacts. As described in Section 2.4, we tried to confirm self-reported enrollment by visiting the schools the girls reported to be enrolled in and asking their teachers about their enrollment statuses. Panel B in Table III reports the same information as Panel A, but as reported by the teachers.

The evidence in Panel B reverses the finding on the relative effectiveness of CCTs and UCTs in preventing dropouts. First, we note that the dropout rates in the control group are higher by approximately 5-6 percentage points (pp), consistent with over-reporting by the core respondents. Furthermore, while dropout rates are still lower in both treatment arms than the control group, the impacts in the CCT arm are significantly larger than the UCT arm and the difference between the two treatment arms in terms of total number of terms enrolled during the two-year intervention is significant at the 95% confidence level (p-value=0.011). Finally, the impact of the CCT intervention seems to have persisted after the cash payments stopped at the end of 2009, while the enrollment rate in the UCT arm is identical to that in the control group during Term 1 of 2010.[36]

Given the divergent results using self-reports and teacher-reports, which set of findings should we believe? Baird and Özler (2010) use administrative records from the CCT program to establish that the school ledgers collected independently in Round 3 provide a reliable source to measure attendance. Using attendance information from these school ledgers here as our benchmark, we can examine the extent of misreporting by the students and the teachers. Table IV presents this evidence. In column (1), we see that 17.0% of the girls in the control group who reported being enrolled in school during Term 2 of 2009 were found to have never attended school during that period according to the school ledgers. This likelihood to over-report enrollment is reduced by more than 50% (9.3 pp) in the CCT arm, but is identical in the UCT arm. Column (2) shows that not only is over-reporting substantially reduced when the information comes from the teachers (5.2% in the control group), but also that the differential misreporting disappears.

The analysis above explains the divergent findings in Table III: girls in both the control group and the UCT arm are significantly more likely than the CCT arm to report being enrolled in school when in fact they are not. Thus, self-reported data attenuate program impacts for the CCT arm and give the impression that UCTs outperform CCTs in reducing dropout rates. Teacher

[36] In Appendix C, we conduct a battery of exhaustive robustness checks and conclude that the findings presented in Table III are robust to varying the sub-samples analyzed and the rules by which missing data for teacher-reported enrollment are treated.

reports substantially reduce the bias caused by this differential misreporting and reveal the true program impacts. As we will see below, the evidence regarding program impacts on school attendance and test scores are also consistent with the finding that the CCT arm was more effective in reducing school dropout than the UCT arm.

*Attendance (intensive margin)*

We now turn to examining the intensity of attendance for those enrolled in school in 2009 and the first term of 2010. The school ledgers from Round 3 provide term by term information on the number of days the students were present for each day the school was in session.[37] Table V indicates that the attendance rate in the control group among those enrolled in school ranges from a high of 85% in Term 2, 2009 to a low of 69% in Term 3, 2009. Attendance on the intensive margin is uniformly higher in the CCT arm than the control group. The overall attendance rate for 2009 is 8.0 pp higher in the CCT arm than the control group, which translates into approximately four school days per term or more than ten school days over the entire 2009 school year. In the UCT arm, impact estimates are mostly positive, but none of them are statistically significant. Program impacts in the CCT arm are higher than the UCT arm during Term 1 in both 2009 (13.9 pp vs. 6.3 pp; p-value=0.129) and 2010 (9.2 pp vs. -3.8 pp; p-value=0.010). Term 1 coincides with the lean season in Malawi, when food is scarcest and the number of malaria cases reaches its peak.[38] Thus, the condition to attend school regularly seems most effective in keeping attendance rates high when households need cash the most.[39]

[37] Appendix Table B.2 confirms balance in this sub-sample of girls for whom attendance data from ledgers are available.

[38] In 2001, the prevalence of malaria parasitaemia among non-pregnant females, ages 15-19, was 24% (Dzinjalamala 2009). The same figure was 47% in school children. Malaria is a frequent cause of absenteeism in school, resulting in poor scholastic performance on the part of the student.

[39] CCT households may have also taken additional measures to minimize school absence by having the girls sleep under bed nets: the share of girls who reported sleeping under a bed net the previous night was 10 pp higher (p-value=0.049) in the CCT arm than the UCT arm in Round 3.

*Test Scores*

In Table VI, we present program impacts on tests of English reading comprehension, mathematics, and cognitive ability, which were administered to all study participants at their homes.[40] We see across-the-board improvements in test scores in the CCT arm, while no significant improvement can be detected in the UCT arm. The 0.140 SD improvement (p-value=0.010) in English reading comprehension in the CCT arm is significantly higher than the program impact in the UCT arm at the 90% confidence level (p-value=0.069). The CCT arm also has a 0.114 SD advantage over the UCT arm in the TIMMS math score, but this difference is not statistically significant.[41] Finally, in terms of cognitive ability, measured by Raven's colored progressive matrices, we see improvements of 0.174 and 0.136 SD in the CCT and UCT arms, respectively, compared with the control group. However, while the improvement in the CCT arm is statistically significant at the 99% confidence level, the impact estimate for the UCT arm is noisy and insignificant.

Summarizing the program impacts on schooling outcomes in the two treatment arms, we find that the CCT arm had significant gains in enrollment on the extensive margin, some improvement in attendance on the intensive margin, and modest gains in achievement in tests of English reading comprehension, mathematics, and cognitive skills.[42] Girls in the UCT arm were also significantly more likely to be enrolled in school compared with the control group, but there was no detectable improvement in their intensity of school attendance or their test scores. The increase in

---

[40] We presented evidence in Table I that there is no differential attrition between the CCT and UCT arms. Moreover, in the sample used for analysis in this paper, i.e. those who were successfully interviewed in all three rounds, only 30 girls have missing test scores (22 in the control group, 5 in the CCT arm, and 3 in the UCT arm). Even if all the controls were assigned the highest score for the English test and all treatment girls were assigned the lowest score, the findings on treatment impacts would not change.

[41] TIMMS stands for Trends in Mathematics and Science Study, which is a cycle of internationally comparative assessments in mathematics and science carried out at the fourth and eighth grades every four years. We borrowed five mathematics questions from the 2007 TIMMS (four fourth-grade and one eighth-grade question) and incorporated them into our independently developed mathematics test.

[42] The improvements in test scores is different than what has been previously reported in evaluations of other CCT programs. Behrman, Parker and Todd (2009) and Filmer and Schady (2009) find no impacts of CCTs on tests of mathematics and language in Mexico and Cambodia, respectively.

enrollment (measured by the total number of terms enrolled in school during the two-year program) in the UCT arm was less than half of that achieved in the CCT arm. We conclude that CCTs outperformed UCTs in terms of improvements in schooling outcomes.

**3.3.2. Marriage and Pregnancy**

We now turn to an analysis of the prevalence of marriage and pregnancy in Rounds 2 and 3. Column (1) of Table VII shows that, by Round 2, 4.3% of the initially never-married sample was married in the control group. Marriage rates were unchanged in the CCT arm, but significantly lower in the UCT arm. By Round 3, the prevalence of marriage rose to 18.0% in the control group with an insignificant reduction of 1.2 pp in the CCT arm and a significant reduction of 7.9 pp (44%) in the UCT arm (column (2)). The differences in program impacts between the two treatment arms in Rounds 2 and 3 are both statistically significant at the 95% confidence level. Columns (3 & 4) show that while the program had little effect on the prevalence of pregnancy in either treatment arm after one year, there is a large decline in the UCT arm by Round 3. While roughly a quarter of the control group and the CCT arm had been 'ever pregnant' by the end of the experiment, this likelihood was reduced by 6.7 pp (or 27%) in the UCT arm – significant at the 99% confidence level. The difference in program impacts between the two treatment arms is also significant at the 99% confidence level by Round 3.

How do we reconcile the fact that the CCT arm led to schooling increases with the fact that the sharp decreases in marriage and pregnancy are found only among the UCTs? Existing evidence from sub-Saharan Africa suggests that reducing school dropout should lead to declines in teen marriage and pregnancy rates (Duflo, Dupas, and Kremer 2010; Ozier 2011; Ferré 2009, Osili and Long 2008). There could also be an income effect on marriage if girls are marrying early due to poor economic circumstances (Field and Ambrus 2008), and on fertility if monthly transfers allow girls to reduce transactional sex with 'sugar daddies' (Dupas 2010; Duflo, Dupas, and Kremer 2010).

To understand the differential impacts of CCTs and UCTs on any outcome other than schooling itself, it is useful to consider three latent strata of baseline schoolgirls. In the first stratum, UCT offers would be sufficient to keep baseline schoolgirls enrolled in school (*UCT compliers*).[43] This group would receive transfers under both treatments, so program impacts could only differ due to behavior changes on the intensive margin arising from the conditionality. In the second stratum are girls who would be enrolled in school in the CCT arm, but not in the UCT arm (*CCT compliers*). In this group, which would also receive transfers under both treatments, conditionality would generate a differential impact on enrollment by lowering the opportunity cost of schooling in the CCT arm. Finally, in the third stratum are girls who would drop out of school under either treatment arm and would receive transfers only under the UCT (*non-compliers*).

The differential impacts of CCTs and UCTs on outcomes other than schooling, such as marriage and pregnancy, will depend on the relative sizes of these three strata and the extent to which the relevant outcomes are driven by schooling and income. We would expect CCTs to be more effective for outcomes such as test scores, which would likely improve with schooling (and the incentives associated with schooling) and are not likely to be affected by increased income among dropouts. Conversely, UCTs may be more effective for other outcomes if there is a large group of *non-compliers*, a strong income effect among them, and small incentive effects among those enrolled in school. In the analysis below, we apply this simple framework to our data and show that, in our setting, marriage is a perfect example of the latter case.

A simple way of presenting the sizes of these three strata and the approximate magnitude of the effects within each of them is provided in Table VIII, which gives marriage rates by treatment

[43] This group includes those who would be enrolled in school in the absence of either treatment. We assume that UCT offers have no perverse effects on schooling, meaning that an offer of UCTs would not cause a girl to drop out of school if she would be enrolled otherwise. We make a similar 'no defiers' assumption with respect to the conditionality: if a girl would be enrolled in school under a UCT, then she would also be enrolled under a CCT offering the same transfer amount.

status and teacher-reported enrollment status in Round 3.[44] For each treatment arm, the top row summarizes the marriage rates by follow-up enrollment status, while the bottom row shows the raw follow-up enrollment rate in that group.[45]

The first thing to note in Table VIII is that, on average, only 1.1% of those still enrolled in school are married (column (1)). This confirms the notion that marriage and school enrollment are practically mutually exclusive in this setting and implies that CCTs cannot generate a differential impact on marriage among *UCT compliers*.[46] Second, a comparison of marriage rates in the CCT arm and the control group among those not enrolled in school confirms the prediction that CCTs are unlikely to have an impact among *non-compliers* as girls in this group stop receiving transfers when they drop out of school (column (2)). Approximately half of the girls not enrolled in school were married in both the control group and the CCT arm in Round 3.

These two simple observations imply that the only way in which CCTs can reduce marriage rates at follow-up is by averting dropouts. Column (1) in Table IX shows that school dropout was 5.8 pp lower in the CCT arm than the control group. Column (2) shows that the reduction in the likelihood of being 'ever married' in the CCT arm is 2.6 pp in this sub-sample of girls for whom data on teacher-reported enrollment are available in Round 3. This small CCT effect on marriage is exactly what we would expect to find given the roughly 50 pp difference in marriage rates between girls enrolled in school and those not enrolled (-0.058*0.5 = -0.029). We conclude that the program impact on dropout in the CCT arm was too small to cause a significant reduction in the likelihood of being married.

---

44 The three strata are defined over latent choices, so Table VIII presents instead compliance and marriage rates as actually observed in the control group and each treatment arm in Round 3. This gives us the unadjusted size of each stratum (*UCT compliers* are 61%, *CCT compliers* are 8%, and *non-compliers* are 31%), but does not tells us which individuals belong to which stratum.

45 For expositional purposes, it is easiest to present marriage and enrollment rates at the time of Round 3 data collection, i.e. during Term 1 of 2010. The findings from the analysis that follows remain qualitatively the same if we use enrollment data for any school term in 2008 or 2009, utilize marriage rates reported in Round 2, or replace marriage with pregnancy.

46 Ozier (2011), summarizing the literature on schooling and fertility, states that they are in practice nearly mutually exclusive in Kenya and argues that this is true in many other contexts.

The UCT arm presents a contrasting case, in which the program impacts on marriage and fertility arise only through changes in behavior among those who dropped out of school. Table VIII shows that the raw marriage rate in the UCT arm among those not enrolled in school is approximately 50% lower than the CCT arm and the control group. In Table IX, we see that the UCT impact on marriage was -8.8 pp, while there was no impact on dropping out of school. Hence, there can be no 'schooling' channel operating on marriage and the decline in the UCT arm is entirely due to the effect of cash transfers among those who dropped out of school. The final two columns in Table IX present regressions of being 'ever married' for those enrolled in school and those who dropped out, respectively, and shows that the marriage rate among those who dropped out is significantly lower in the UCT arm compared to both the control group and the CCT arm.[47]

The evidence presented here is consistent with the literature that posits female schooling as a means of delaying marriage and fertility, but the product of the treatment effect on dropping out in the CCT arm (5.8 pp) and the sizeable effect of school enrollment on marriage (reduced by nearly 50 pp) is too small to be detectable given our sample size.[48] The novel finding here is the large impact of unconditional cash transfers on delaying marriage and pregnancy among adolescent girls who dropped out of school. The effect of UCTs on marriage among this group (reduced by more than 20 pp) is smaller than the schooling effect on marriage, but experienced in a large enough sub-group (40% of the UCTs had dropped out by Round 3) to be strongly significant in the whole UCT arm. The 'schooling' channel on marriage operates through the dropouts averted by the CCT arm, while the 'income' effect on marriage operates through those who dropped out of school; and the latter group is substantially larger than the former in this experiment.

---

[47] It should be noted that these are not proper sub-samples on which experimental program impacts can be estimated as they are constructed using enrollment status at follow-up, which is endogenous to treatment. We only present these regressions to give a sense of the magnitude and the statistical significance of the raw differences in Table VIII.

[48] To confirm the potential importance of the schooling channel, note that Baird et al. (2010) show that CCTs had a significant impact on reducing marriage and pregnancy rates among baseline dropouts by 11.3 pp and 5.1 pp, respectively. In this sample the CCT led to a very large increase in self-reported re-enrollment in comparison to the control group (61% vs. 17%), and hence the impact through the 'schooling' channel is detectable.

### 3.3.3. Heterogeneity of program impacts

We now turn to an examination of impact heterogeneity, using only covariates that were used to stratify our initial study design (age) or were directly randomized (the transfer amounts made to the girls and their parents).[49] Table X examines heterogeneity using an indicator for being sixteen years or older.[50] With respect to enrollment, program impacts do not vary significantly by age group: CCTs outperform UCTs in raising enrollment among early adolescents as well as older teenagers. However, when it comes to the other three main outcomes examined in this study, we see that the advantage in English test scores in the CCT arm among early adolescents disappears among girls sixteen or older at baseline, while the advantage the UCT arm has in preventing marriages and pregnancies is substantially larger among older teenagers. For each of these three outcomes, the difference in the coefficients for the interaction terms between the CCT and UCT arms is large and statistically significant. Our results indicate that, among early adolescents, there is a clear trade-off between a CCT and a UCT program in terms of improved schooling outcomes versus a reduction in teenage pregnancies. Among older teenage girls, however, this trade-off largely disappears and UCTs become relatively more attractive.[51]

We next examine the heterogeneity in program impacts by two important program features: the identity of the transfer recipient within the household and transfer size. Bursztyn and Coffman (2010) argue that policies designed to promote school attendance might be more effective if they

[49] The study sample was stratified by age group at baseline as program impacts on enrollment, marriage, and pregnancy can reasonably vary along this dimension. Transfer amounts made separately to girls and their parents were also randomized to assess whether transfer size or the identity of the recipient within the household – two key policy parameters in cash transfer programs – affected the outcomes. We limit ourselves to examining heterogeneity of program impacts only along these baseline characteristics that were defined ex-ante in order to avoid data-mining.

[50] The legal age of marriage stood at 16 in Malawi by late 2009 (Nyasa Times September 22, 2009) and many girls aged 16 or older are attending primary or secondary school. In our study sample, 22% of students eligible to attend primary school at baseline were aged 16 or older. This share increases to 56% for those eligible to attend secondary school at baseline.

[51] We have also examined heterogeneity of program impacts using a household asset index to imitate the means-tested targeting schemes like the ones in Brazil or Mexico. In Appendix Table D.1, we show that enrollment impacts in the UCT arm may improve somewhat under such a targeting scheme, but otherwise program impacts would not be significantly altered compared to a universal cash transfer program for never-married adolescent girls.

target the child instead of focusing on parents because sub-optimal school attendance may be due to a parent-child conflict, where the parents cannot enforce their desire that their children attend school. Similarly, World Bank (2009) argues that "…the key parameter in setting benefit levels is the size of the elasticity of the relevant outcomes to the benefit level." However, random variation in these design parameters is rarely observed in cash transfer programs around the world. In this experiment, separate transfers were offered to girls and their parents (or guardians), the size of each of which were randomly determined.

We analyze the impact of transfer amounts by estimating the following linear probability model:

(2) $$Y_i = I_i^C\delta^C + I_i^U\delta^U + H_i^C\phi^C + H_i^U\phi^U + T_i^C\gamma^C + T_i^U\gamma^U + X_i\beta + \varepsilon_i$$

$T_i^C$ and $T_i^U$ are again dummy variables for the CCT and UCT offers, $I_i^C, I_i^U$ and $H_i^C, H_i^U$ give the individual and household transfer amounts for each treatment, defined in differences from the lowest amount offered in the treatment arms ($1 for the individual transfer, $4 for the household transfer). The estimates of $\delta$ and $\phi$ thus give the marginal effect of increasing individual and household amounts by $1 under each treatment arm, while the estimates of $\gamma$ measure the impact of each treatment at the lowest total transfer amount. The standard errors $\varepsilon_i$ are again clustered at the EA level.

Table XI presents the heterogeneity in program impacts by transfer amounts. Column (1) shows program impacts on enrollment: the minimum total transfer amount offered to the household ($5/month) seems to be responsible for the entire program impact on the 'total number of terms enrolled' in the CCT arm. Additional transfers offered to either party make little difference. In the UCT arm, however, the effect of the minimum transfer offer is small and insignificant, but enrollment increases 0.081 terms with each additional dollar offered to the parents over and above

\$4/month.[52] The analysis implies that giving an additional \$5/month in transfers to the parents would barely allow the UCT arm to reach the level of enrollment attained by the minimum transfer amounts in the CCT arm. As the marginal administrative cost of a CCT program is likely to be only a small share of each dollar transferred to beneficiary households, a CCT program offering \$5/month would be substantially more cost-effective in increasing enrollment than a UCT program offering the same amount.

For English test scores, the results are similar for the CCT arm: there is no indication that additional amounts to the girls or their parents would improve test scores over and above the minimum monthly transfer. Here, unlike enrollment, the coefficient in the UCT arm is similar to the CCT arm at the minimum transfer amounts, although insignificant. For marriage, there is no treatment impact at the minimum transfer amounts in the UCT arm, but each additional dollar offered to the parents of a girl reduces her likelihood of getting married by Round 3 by 1.7 pp.[53] The marginal effect of a dollar offered to parents in delaying marriages is significantly higher in the UCT arm than that in the CCT arm (p-value=0.069). The large negative and statistically significant gradient between transfer size and likelihood of marriage at follow-up is consistent with parents marrying their daughters early due to economic considerations. Finally, the minimum amounts transferred in the UCT arm seem to be responsible for almost the entire program effect on preventing pregnancies in this group.

In summary, increasing transfer amounts or varying the recipient within the household has no effect on any of the outcomes examined in this paper in the CCT arm; this contract variation

[52] The difference between the CCT and UCT impacts on enrollment at the minimum transfer amount (\$5/month total to the household) is significant at the 95% confidence level.
[53] Similarly, each additional dollar offered to the girl in the UCT arm reduces her chances of being 'ever married' at follow-up by 1.6 pp (p-value=0.148).

simply does not seem to matter for CCTs.[54] In contrast, we find that outcomes vary with increased transfer offers to the parents in the UCT arm: enrollment rates increase and the incidence of marriage declines as parents are offered more money, but performance in test scores seems to suffer. Still, however, replacing a CCT program that offers the minimum transfer amounts of $1 to the girl and $4 to her parents with a UCT program that offers the parents larger transfer amounts would not be cost-effective in improving schooling outcomes, but it would reduce marriage rates among teenage girls. Furthermore, we find no evidence that increasing the amount of transfers made directly to the girl rather than her parents would be effective in improving any of the outcomes studied here.

Before we move to the concluding section, we briefly discuss the robustness of our findings with respect to two issues. First, treatment and conditionality status were assigned at the EA level in this experiment. Due to the proximity of EAs to each other, it is possible that the intermingling of students in the two treatment arms led to a change in behavior in the outcomes of interest, thus biasing our estimates of the marginal impact of the conditionality. Second, as described in Section 2, the offers in the CCT arm included a promise to pay secondary school fees directly to the schools upon confirmation of enrollment by the program administrators. To make the average transfer offers in the UCT arm equal to that in the CCT arm, the average school fee amount was added to the cash transfer offers of girls in the UCT arm who were eligible to attend secondary school at the beginning of the program. In Appendix E, we discuss each of these issues in some detail and conclude that it is unlikely that our estimates of differential program impacts are influenced by either spillovers due to the proximity of girls with discordant treatment statuses or by the manner in which school fee compensation was handled.

---

[54] This finding is consistent with the evidence presented in Filmer and Schady (2010), which found substantial impacts on enrollment of a modest CCT offer in Cambodia, but no additional enrollment gains from a larger transfer offer. On the other hand, Barrera-Osorio, Linden, and Urquiola (2007) find enrollment responses of students to be sensitive to the size of the implied subsidies in Bogota's *Gratuidad* school fee reduction initiative.

## 4. Concluding discussion and policy implications

This paper presented experimental evidence on the relative effectiveness of conditional and unconditional cash transfer programs. The analysis focused on two sets of outcomes that are of central importance to the long-term prospects of school-age girls: schooling and human capital formation on the one hand; marriage and fertility on the other. The results show that CCTs increased enrollment rates and improved regular attendance for those in school, both of which likely contributed to a modest but significant improvement in English test scores over the UCT arm. Teenage pregnancy and marriage rates, on the other hand, were substantially lower in the UCT than the CCT arm.

The results on school enrollment differ from previous studies that considered the relative effectiveness of UCTs vs. CCTs, although the difference is a matter of degree rather than direction. Analyses of Bolsa Escola in Brazil and PROGRESA in Mexico found that UCTs would have little, if any, impact on school enrollment – implying that almost all of the impacts of these programs were due to the schooling condition (Bourguignon, Ferreira, and Leite 2003; Todd and Wolpin 2006). In the Malawi experiment, on the other hand, we find a modest impact on enrollment in the UCT arm, which is 43% as large as the impact in the CCT arm. Our study thus confirms that conditions attached to cash transfer programs are effective in increasing enrollment, but the size of this effect is likely to be smaller than suggested by earlier studies, at least for poorer countries like Malawi.

Not only is school enrollment significantly improved in the CCT arm over the UCT arm, but the evidence presented shows that CCTs are more cost-effective in raising enrollment than UCTs in this context. To achieve the same enrollment gain obtained from a \$5/month total transfer in the CCT arm, a transfer of more than \$10 to the parents in the UCT arm is needed. This difference is much larger than the additional cost of administering a CCT program – possibly by an order of magnitude.[55] Furthermore, the average number of payments in the CCT arm was approximately 14.1 (out of a possible total of 20 over two years), compared with 17.9 in the UCT arm, implying that the actual amount of transfers made per person was 19% lower in the CCT arm over the two-year intervention. Savings of this magnitude due to non-compliance with the schooling condition would more than make up for the additional administrative cost of monitoring in most programs.

While CCTs were more cost-effective than UCTs in increasing school enrollment and attendance, they had little effect on reducing the likelihood of teenage pregnancies or marriages. By contrast, UCT offers were very effective in delaying marriage and childbearing – by 44% and 27%, respectively, after two years. These impacts in the UCT arm were experienced almost entirely among those who dropped out of school after the start of the two-year intervention, while the likelihood of marriage and pregnancy were negligible among those who stayed in school regardless of treatment status.

When we examine the random variation in transfer amounts to girls and their parents, we find that this contract variation did not alter outcomes in the CCT arm. This finding is encouraging for policy-makers as it implies that the smallest transfer amounts ($4/month to the parents and $1/month to the school-age girl) offered in this experiment were sufficient to attain the average schooling impacts observed under the CCT arm. Furthermore, our results suggest that reallocating some of the transfers from the parents to the girls would not improve program impacts under either treatment. Hence, while the idea of making part of the transfers directly to adolescent girls may be attractive on its face, we find no evidence that it would be effective.[56]

---

[55] World Bank (2009) cites Grosh et al. (2008) to report that the median administrative cost is 8% of total program costs in ten CCT programs for which administrative cost data were compiled by that study. Given that monitoring compliance with the schooling condition is only part of the administrative costs, the additional cost would be a few cents on the dollar. However, monitoring and enforcement of conditionality in many of these programs may not have been as thorough as they were in our experiment, which could reduce the relative effectiveness of CCTs compared with UCTs.

[56] The readers should note, however, that there was no cell in the experiment in which the transfer to the girl (or her parents) was equal to zero. It is possible that even a $1/month transfer made directly to the girls was more effective than

The experiment highlights the underlying causes of school dropout, marriage, and pregnancy in the Malawian context. First, the fact that the UCT has some effect on school enrollment indicates that poverty is a cause of school dropout in this population, and that poor parents will invest at least some of the additional funds from a positive income shock towards the education of their daughters. The reduction in marriage and pregnancy rates in the UCT arm, as well as the responsiveness of marriage to unconditional transfer amounts, seems consistent with the idea that adolescent girls who drop out of school undergo a rapid transition into adulthood that is also strongly influenced by economic circumstances.[57]

The informal framework we introduced in Section 3.3.2 is helpful in considering the generalizability of differential program impacts on marriage and pregnancy found in this study. A necessary condition for UCTs to delay marriage and pregnancy more effectively than CCTs is the presence of a large group of *non-compliers* (relative to the group of *CCT compliers*), a finding that is confirmed in most evaluations of CCT programs around the world.[58] Given a large stratum of *non-compliers*, UCTs may be more effective in delaying marriage and pregnancy if there are significant beneficial income effects among this group. In Southern and Eastern Africa, where these outcomes are more or less mutually exclusive with schooling and decisions regarding sexual behavior and marriage among adolescent girls are influenced by poverty, UCTs may indeed be more effective than schooling CCTs in reducing teenage marriage and pregnancy. In countries like Bangladesh, where

---

adding that small amount to the transfers made to the parents. Our results are clear that marginal effect of reallocating any amount above the $1/month given to the girl would not improve outcomes.

[57] In our study sample, approximately 25% of the young women who were sexually active at baseline reported that they started their sexual relationships because they "needed his assistance" or "wanted gifts/money."

[58] Most studies examining school dropout in CCT programs find that dropout rates are reduced by less than 50%, implying that the stratum of *non-compliers* is usually larger than the stratum of *CCT compliers*. For example, Behrman, Sengupta, and Todd (2005) report that, among 15 year-olds enrolled in school before the start of PROGRESA, the dropout rate in the treatment group was 31.3%, while the program impact on reducing dropouts was 6.4 pp. These numbers are remarkably similar to those presented in Table IX. Even in countries like Cambodia, where the program impacts on enrollment were very large in absolute terms at approximately 25 pp, the reduction in the dropout rate was less than 50% as the enrollment rate in the control group was 44% (Filmer and Schady 2010). It should be noted that program impacts on dropout rates reported in studies of CCT programs are upper bounds for the size of the *CCT complier* group as there were no counterfactuals for a UCT intervention.

dropout and marriage rates among adolescent girls are also high but, unlike sub-Saharan Africa, dowry payments are made from the bride's family to the groom's, UCTs may have no effect, or perhaps even the opposite effect, on the timing of marriage. In such settings, CCTs for schooling or for staying unmarried may be more effective than UCTs.[59]

Our framework also allows us to consider the relative merits of CCT and UCT programs on outcomes other than schooling, marriage, and pregnancy. CCTs are likely to be more effective in improving outcomes that may be strongly affected by compliance with the conditions, such as test scores. UCTs may be preferred if there are many *non-compliers* who might experience strong and socially beneficial effects from regular income support. If *non-compliers* can be thought of as a vulnerable group in a given context, UCTs may deserve careful consideration given the possible trade-offs indicated in this study. While similar experiments would be useful in determining the extent to which our findings generalize to other contexts, in the absence of such an experiment, policymakers can use evidence from previous evaluations of CCT programs to estimate the sizes of the relevant strata and use existing household survey data to predict the relative effects of schooling and income on outcomes of interest within each stratum.

Whatever the overall merits of the two transfer schemes, this study found evidence that schooling CCTs are a much more cost-effective means of reducing dropouts than are UCTs. However, in the absence of a market failure, such a distortion is inefficient.[60] Policy-makers planning to implement a CCT program should clearly articulate the market failures behind this paternalistic motivation and, if possible, provide evidence of these externalities – private or social. For example, we have not found any evidence of incomplete altruism – i.e. a conflict of interest between the girls

[59] Such an intervention is being currently evaluated in Bangladesh, where girls between 15 and 17 years old living in treatment areas, are offered cooking oil on the condition that they remain unmarried. The amount of oil offered is designed to be larger than the cost of delaying marriage, which comes in the form of increased dowry payments from the bride's family to the groom's: http://www.povertyactionlab.org/evaluation/empowering-girls-rural-bangladesh.

[60] In this context, credit market failures would not be a justification for CCTs as UCTs would be sufficient to relax credit constraints. Another market failure, such as a lack of information, hyperbolic discounting, or positive spillover effects, is necessary to justify preferring CCTs over UCTs.

and their parents with respect to her education, which is sometimes mentioned as a justification of a preference for CCTs over UCTs.

CCT programs create incentives for individuals to change their behaviors by denying transfers to those who fail to satisfy the conditions. However, at least some of these individuals come from vulnerable households and are also in need of income support. Our findings suggest that UCTs to such households can improve important outcomes even though they are not as successful in improving schooling outcomes as CCTs. This study makes clear that while CCT programs may be more effective than UCTs in obtaining the desired behavior change, they can also undermine the social protection dimension of cash transfer programs.

GEORGE WASHINGTON UNIVERSITY

UNIVERSITY OF CALIFORNIA, SAN DIEGO

THE WORLD BANK

## References


Adato, Michelle, and Lucy Bassett, "Social Protection to Support Vulnerable Children and Families: the potential of cash transfers to protect education, health and nutrition," *AIDS Care*, 21(2009), 60-75.

Akinbami, Lara J., Kenneth C. Schoendorf, and John L. Kiely, "Risk of Preterm Birth in Multiparous Teenagers," *Archives of Pediatrics and Adolescent Medicine*, 154(2000), 1101-1107.

Ashcraft, Adam, and Kevin Lang, "The Consequences of Teenage Childbearing," NBER Working Paper 12485. National Bureau of Economic Research, Cambridge, MA, 2006.

Baird, Sarah, Ephraim Chirwa, Craig McIntosh, and Berk Özler, "The Short-Term Impacts of a Schooling Conditional Cash Transfer Program on the Sexual Behavior of Young Women," *Health Economics*, 19(2010), S1: 55-68.

Baird, Sarah, and Berk Özler, "Examining the Reliability of Self-Reported Data on School Participation," *unpublished manuscript*, 2010.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez-Calle, "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia," *American Economic Journal: Applied Economics*, forthcoming, 2010.

Barrera-Osorio, Felipe, Leigh L. Linden, and Miguel Urquiola, "The Effects of User Fee Reductions on Enrollment. Evidence from a Quasi-Experiment," *unpublished manuscript*, 2007: http://www.columbia.edu/~ll2240/Gratuidad%20Draft%202007-01.pdf.

Behrman, Jere R., Susan W. Parker, and Petra E. Todd, "Medium-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico" in *Poverty, Inequality and Policy in Latin America*, eds. Stephan Klasen and Felicitas Nowak-Lehman, 219-70 Cambridge, MA, MIT Press, 2009.

Behrman, Jere R., Piyali Sengupta, and Petra Todd, "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico," *Economic Development and Cultural Change*, 54 (2005), 237-275.

Blau, Francine, and Lawrence Kahn, "Do Cognitive Test Scores Explain Higher U.S. Wage Inequality?" *Review of Economics and Statistics,* 87(2005), pp. 184-193.

Bloom, David E., David Canning, Günther Fink, and Jocelyn E. Finlay, "Fertility, Female Labor Force Participation, and the Demographic Dividend," *Journal of Economic Growth*, 14(2009), 79-101.

Bourguignon, François, Francisco H.G. Ferreira, and Phillippe G. Leite, "Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Brazil's Bolsa Escola Program," *The World Bank Economic Review,* 17(2003), 229-254.

Bruhn, Miriam, and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1(2008), 200-232.

Bursztyn, Leonardo, and Lucas Coffman, "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian *Favelas," unpublished manuscript*. http://www.anderson.ucla.edu/faculty/leonardo.bursztyn/Schooling_Decision_10_18_10.pdf, 2010.

Case, Anne, Victoria Hosegood, and Frances Lund, "The reach and impact of Child Support Grants: Evidence from KwaZulu-Natal," *Development Southern Africa*, 22(2005), 467-482.

Cunha, Jesse M., "Testing Paternalism: Cash vs. In-kind Transfers in Rural Mexico," *unpublished manuscript*. http://www.stanford.edu/~jcunha/Cunha_Testing_Paternalism.pdf, 2010.

Dahl, Gordon B., "Early Teen Marriage and Future Poverty," *Demography*, 47(2010), 689-718.

de Brauw, Alan, and John Hoddinott, "Must Conditional Cash Transfer Programs be Conditioned to be Effective? The Impact of Conditioning Transfers on School Enrollment in Mexico," *Journal of Development Economics* (2010), doi:10.1016/j.jdeveco.2010.08.014.

de Janvry, Alain, Frederico Finan, Elisabeth Sadoulet, and Renos Vakis, "Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?" *Journal of Development Economics*, 79(2006), 349-373.

de Janvry, Alain, and Elisabeth Sadoulet, "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality," *The World Bank Economic Review*, 20(2006), 1-29.

Duflo, Esther, "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa," *The World Bank Economic Review*, 17(2003), 1-25.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Education and Fertility: Experimental Evidence from Kenya," *unpublished manuscript*, 2010:

http://econ.duke.edu/uploads/assets/Workshop%20Papers/Education&Fertility_feb2010.pdf.

Dupas, Pascaline, “"Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, 3(2011), 1–34.

Dzinjalamala, Fraction, "Malaria," in *The Epidemiology of Malawi*, eds. (Geubbels and Bowie), College of Medicine, University of Malawi, 2009:

http://www.medcol.mw/commhealth/publications/epi%20book/epidemiology%20book.htm

Edmonds, Eric V., "Child Labor and Schooling Responses to Anticipated Income in South Africa," *Journal of Development Economics*, 81(2006), 386-414.

Edmonds, Eric V., and Norbert Schady, "Poverty Alleviation and Child Labor," NBER Working Paper 15345. National Bureau of Economic Research, Cambridge, MA, 2009.

Ferré, Celine, "Age at First Child: Does Education Delay Fertility Timing ? The case of Kenya," World Bank Policy Research Working Paper 4833, 2009.

Field, Erica, and Attila Ambrus, "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh," *Journal of Political Economy*, 116(2008), 881-930.

Filmer, Deon, and Norbert Schady, "Does More Cash in Conditional Cash Transfer Programs Always Lead to Larger Impacts on School Attendance?" *Journal of Development Economics* (2010), doi:10.1016/j.jdeveco.2010.05.006.

___________, "School Enrollment, Selection and Test Scores," World Bank Policy Research Working Paper 4998, 2009.

Fletcher, Jason M., and Barbara L. Wolfe, "Education and Labor Market Consequences of Teenage Childbearing: Evidence Using the Timing of Pregnancy Outcomes and Community Fixed Effects," NBER Working Paper 13847. National Bureau of Economic Research, Cambridge, MA, 2008.

Fraser, Alison M., John E. Brockert, and R.H. Ward, "Association of Young Maternal Age with Adverse Reproductive Outcomes," *The New England Journal of Medicine*, 332(1995), 1113-1117.

Goldin, Claudia, and Lawrence F. Katz, "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 110(2002), 730-770.

Hanlon, Joseph, Armando Barrientos, and David Hulme, *Just Give Money to the Poor: The Development Revolution from the South,* Kumarian Press: USA, *2010*.

Hanushek, Eric V., and Ludger Woessmann, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," NBER Working Paper 14633. National Bureau of Economic Research, Cambridge, MA, 2009

Horgan, Richard P., and Louise C. Kenny, "Management of Teenage Pregnancy," *The Obstetrician & Gynaecologist*, 9(2007), 153–158.

Hotz, V. Joseph, Susan W. McElroy, and Seth G. Sanders, "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Very Natural Experiment" *Journal of Human Resources* 40(2005), 683–715.

Jensen, Robert, and Rebecca Thornton, "Early female marriage in the developing world," *Gender and Development*, 11(2003), 9-19.

Kremer, Michael, Edward Miguel, and Rebecca Thornton, "Incentives to Learn," *The Review of Economics and Statistics*, 91(2009), 437-456.

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

Macours, Karen, Norbert Schady, and Renos Vakis, "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment," World Bank Policy Research Working Paper, 2008.

Miguel, Edward, and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica,* 72(2004), 159-217.

Miller, Candace, and Maxton Tsoka, "Evaluating the Mchinji Social Cash Transfer Pilot," http://www.unicef.org/socialpolicy/files/REvised_Presentation_Evaluating_the_Mchinji_Social_Cash_Transfer_Pilot_2_July_07.pdf, 2007.

*Nyasa Times*, "Debate over Recommended Marriage Age for Girls Continues in Malawi," September 22, 2009. (http://www.nyasatimes.com/features/debate-over-recommended-marriage-age-for-girls-continues-in-malawi.html)

Osili, Una Okonkwo, and Bridget Terry Long, "Does Female Schooling Reduce Fertility? Evidence from Nigeria," *Journal of Development Economics*, 87(2008), 57-75.

Ozier, Owen, "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis," *unpublished manuscript*, 2011, http://economics.ozier.com/owen/papers/ozier_JMP_20110117.pdf

Paxson, Christina, and Norbert Schady, "Does money matter? The Effects of Cash Transfers on Child Development in Rural Ecuador," *Economic Development and Cultural Change*, 59(2010), 187-229.

Schady, Norbert R., and Maria Caridad Araujo, "Cash Transfers, Conditions, and School Enrollment in Ecuador," *Economía*, 8(2008), 43-70.

Schultz, T. Paul, "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program," *Journal of Development Economics,* 74(2004), 199-250.

Smith, Gordon C. S., and Jill P. Pell, "Teenage Pregnancy and Risk of Adverse Perinatal Outcomes Associated with First and Second Births: Population Based Retrospective Cohort Study," *BMJ*, 323(2001), 1-5.

Todd, Petra E., and Kenneth I. Wolpin, "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96(2006), 1384–1417.

World Bank, *Conditional Cash Transfers: Reducing Present and Future Poverty,* eds: Fiszbein and Schady. World Bank Publications, Washington, DC, USA, 2009.

World Development Indicators Database. 2010. Accessed November, 2010.

Table I: Analysis of Attrition

| | Dependent Variable: | | | | | |
|---|---|---|---|---|---|---|
| | =1 if surveyed in Round 3 | =1 if surveyed in all three Rounds | =1 if took educational tests | =1 if information found in Round 2 school survey | =1 if information found in Round 3 school survey | =1 if legible ledger found |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Conditional treatment | 0.020 | 0.021 | 0.029* | 0.033 | -0.000 | 0.116* |
| | (0.015) | (0.030) | (0.016) | (0.024) | (0.027) | (0.064) |
| Unconditional treatment | 0.021 | 0.030 | 0.035* | -0.029 | 0.014 | 0.061 |
| | (0.019) | (0.024) | (0.020) | (0.053) | (0.028) | (0.077) |
| Mean in the control group | 0.946 | 0.893 | 0.929 | 0.890 | 0.935 | 0.378 |
| Number of observations | 2,284 | 2,284 | 2,284 | 2,284 | 983 | 821 |
| Prob > F(Conditional=Unconditional) | 0.965 | 0.797 | 0.801 | 0.246 | 0.627 | 0.513 |

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. Regressions are restricted to the sub-sample of core respondents who were in school at baseline and sampled to be part of the conditional, unconditional or control sample. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table II: Baseline Means and Balance

| | Mean (s.d) | | | p-value (conditional-unconditional) |
|---|---|---|---|---|
| | Control group | Conditional group | Unconditional group | |
| **Panel A: Household-level variables** | (1) | (2) | (3) | (4) |
| Household size | 6.432 | 6.384 | 6.662 | 0.202 |
| | (2.257) | (2.146) | (2.075) | |
| Asset index | 0.581 | 0.984 | 1.221 | 0.623 |
| | (2.562) | (2.740) | (2.447) | |
| Female-headed household | 0.343 | 0.252** | 0.245** | 0.899 |
| | (0.475) | (0.434) | (0.431) | |
| Mobile phone access | 0.616 | 0.583 | 0.605 | 0.799 |
| | (0.487) | (0.494) | (0.490) | |
| Household transfer amount | N/A | 6.991 | 6.829 | 0.822 |
| | | (2.319) | (2.101) | |
| **Panel B: Individual-level variables** | | | | |
| Age | 15.252 | 14.952* | 15.424 | 0.007*** |
| | (1.903) | (1.827) | (1.923) | |
| Highest grade attended | 7.478 | 7.246 | 7.896** | 0.004*** |
| | (1.634) | (1.598) | (1.604) | |
| Mother alive | 0.842 | 0.802 | 0.836 | 0.360 |
| | (0.365) | (0.399) | (0.371) | |
| Father alive | 0.705 | 0.714 | 0.759 | 0.288 |
| | (0.456) | (0.453) | (0.428) | |
| Never had sex | 0.797 | 0.797 | 0.775 | 0.582 |
| | (0.402) | (0.403) | (0.419) | |
| Ever pregnant | 0.023 | 0.030 | 0.031 | 0.973 |
| | (0.149) | (0.171) | (0.173) | |
| Individual transfer amount | N/A | 3.090 | 3.033 | 0.606 |
| | | (1.431) | (1.451) | |
| Number of observations | 1356 | 470 | 261 | |

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, while in column (4) stars indicate a significant difference between the conditional and unconditional treatment arms. Standard errors are clustered at the EA level to account for the design effect. Means are weighted to make them representative of the target population in the study EAs. The number of unique observations vary slightly across variables due to missing observations, with the total sample size ranging between 2089 and 2084.

Table III: Program Impact on School Enrollment

Panel A: Program impacts on *self-reported* school enrollment

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year 1: 2008 | | | Year 2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 Terms) | Term 1, Post-program |
| Conditional treatment | 0.007 | 0.019* | 0.041** | 0.049*** | 0.056*** | 0.061*** | 0.233*** | 0.005 |
| | (0.011) | (0.011) | (0.017) | (0.017) | (0.018) | (0.019) | (0.070) | (0.025) |
| Unconditional treatment | 0.034*** | 0.051*** | 0.054*** | 0.072*** | 0.095*** | 0.101*** | 0.406*** | 0.074*** |
| | (0.010) | (0.011) | (0.018) | (0.021) | (0.022) | (0.021) | (0.079) | (0.026) |
| Mean in the control group | 0.958 | 0.934 | 0.900 | 0.831 | 0.800 | 0.769 | 5.191 | 0.641 |
| Number of observations | 2,087 | 2,087 | 2,086 | 2,087 | 2,087 | 2,087 | 2,086 | 2,086 |
| Prob > F(Conditional=Unconditional) | 0.006 | 0.012 | 0.460 | 0.299 | 0.102 | 0.098 | 0.038 | 0.028 |

Panel B: Program impacts on *teacher-reported* school enrollment

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.043*** | 0.044*** | 0.061*** | 0.094** | 0.132*** | 0.113*** | 0.535*** | 0.058* |
| | (0.015) | (0.016) | (0.018) | (0.041) | (0.035) | (0.039) | (0.129) | (0.033) |
| Unconditional treatment | 0.020 | 0.038** | 0.018 | 0.027 | 0.059 | 0.033 | 0.231* | 0.001 |
| | (0.015) | (0.017) | (0.023) | (0.038) | (0.037) | (0.039) | (0.136) | (0.036) |
| Mean in the control group | 0.906 | 0.881 | 0.852 | 0.764 | 0.733 | 0.704 | 4.793 | 0.596 |
| Number of observations | 2,023 | 2,023 | 2,023 | 852 | 852 | 852 | 852 | 847 |
| Prob > F(Conditional=Unconditional) | 0.173 | 0.732 | 0.067 | 0.076 | 0.014 | 0.020 | 0.011 | 0.108 |

Notes: The dependent variable in Panel A is whether the core respondent reported being enrolled in school for the relevant year/term. The dependent variable in Panel B is whether the teacher reported the core respondent being enrolled in school for the relevant year/term. Post-program refers to Term 1, 2010, the first term after the program ended. Total terms refers to the total number of terms enrolled during the program. Columns (4)-(8) in Panel B are restricted to the sub-sample of core respondents sampled for the Round 3 school survey who are also both part of the panel data set and part of the school survey panel. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table IV: Analysis of Reporting Bias on Enrollment

| | Dependent Variable: | |
|---|---|---|
| | Core respondents over-reporting | Teachers over-reporting |
| | (1) | (2) |
| Conditional treatment | -0.093* | -0.021 |
| | (0.052) | (0.035) |
| Unconditional treatment | -0.001 | -0.014 |
| | (0.058) | (0.038) |
| Mean in the control group | 0.170 | 0.052 |
| Number of observations | 325 | 325 |
| Prob > F(Conditional=Unconditional) | 0.018 | 0.794 |

Notes: Regressions are restricted to the sub-sample of core respondents who report being enrolled in school during Term 2, 2009, and have legible ledger data. Over-reporting refers to differences between the student (teacher) reports and the ledger. 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 V: Program Impacts on Attendance from School Ledgers

| | Dependent variable: Fraction of days respondent attended school | | | | |
|---|---|---|---|---|---|
| | Term 1, 2009 | Term 2, 2009 | Term 3, 2009 | Overall 2009 | Term 1, 2010 |
| | (1) | (2) | (3) | (4) | (5) |
| Conditional treatment | 0.139*** | 0.014 | 0.169** | 0.080** | 0.092** |
| | (0.045) | (0.033) | (0.085) | (0.035) | (0.041) |
| Unconditional treatment | 0.063 | 0.038 | 0.118 | 0.058 | -0.038 |
| | (0.056) | (0.033) | (0.102) | (0.037) | (0.053) |
| Mean in the control group | 0.778 | 0.849 | 0.688 | 0.810 | 0.801 |
| Number of observations | 284 | 285 | 192 | 319 | 211 |
| Prob > F(Conditional=Unconditional) | 0.129 | 0.334 | 0.358 | 0.436 | 0.010 |

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. The variable "Overall 2009" is defined for all core respondents who have ledger information for any of the three terms and is constructed by dividing the number of days present by the number of days in session for all terms in which there is information. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table VI: Program Impacts on Test Scores

| | Dependent Variable: | | | |
|---|---|---|---|---|
| | English test score (standardized) | TIMMS math score (standardized) | Non-TIMMS math score (standardized) | Cognitive test score (standardized) |
| | (1) | (3) | (2) | (4) |
| Conditional treatment | 0.140*** | 0.120* | 0.086 | 0.174*** |
| | (0.054) | (0.067) | (0.057) | (0.048) |
| Unconditional treatment | -0.030 | 0.006 | 0.063 | 0.136 |
| | (0.084) | (0.098) | (0.087) | (0.119) |
| Number of observations | 2,057 | 2,057 | 2,057 | 2,057 |
| Prob > F(Conditional=Unconditional) | 0.069 | 0.276 | 0.797 | 0.756 |

Notes: The cognitive test score is based on Raven's Colored Progressive Matrices. Math and English reading comprehension tests were developed based on the Malawian school curricula. Five questions (four from the Fourth Grade test and one from the Eight Grade test) from Trends in Mathematics and Science Study (TIMMS) 2007, which is a cycle of internationally comparative assessments in mathematics and science carried out at the fourth and eighth grades every four years, were added to the Math test. All test scores have been standardized to have a mean of zero and a standard deviation of one in the control group. Regressions are OLS models using Round 3 data with robust standard errors clustered at the EA level. All regressions are weighted to make the results 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, an indicator for ever had sex, and whether the respondent participated in the pilot phase of the development of the testing instruments. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table VII: Program Impacts on Marriage and Pregnancy

| | Dependent variable: | | | |
|---|---|---|---|---|
| | =1 if ever married | | =1 if ever pregnant | |
| | Round 2 | Round 3 | Round 2 | Round 3 |
| | (1) | (2) | (3) | (4) |
| Conditional treatment | 0.007 | -0.012 | 0.013 | 0.029 |
| | (0.012) | (0.024) | (0.014) | (0.027) |
| Unconditional treatment | -0.026** | -0.079*** | -0.009 | -0.067*** |
| | (0.012) | (0.022) | (0.017) | (0.024) |
| Mean in the control group | 0.043 | 0.180 | 0.089 | 0.247 |
| Number of observations | 2,087 | 2,084 | 2,086 | 2,087 |
| Prob > F(Conditional=Unconditional) | 0.024 | 0.025 | 0.265 | 0.003 |

Notes: The dependent variables are having been ever married or ever pregnant at the time of the relevant survey. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table VIII: Prevalence of Being 'Ever Married' by School Enrollment Status during Term1, 2010

| | Enrolled | Not enrolled | Total |
|---|---|---|---|
| | (1) | (2) | (3) |
| Control | 1.7% | 46.9% | 19.9% |
| (row %) | (59.8%) | (40.2%) | (100.0%) |
| Conditional treatment | 0.5% | 50.8% | 16.0% |
| (row %) | (69.2%) | (30.8%) | (100.0%) |
| Unconditional treatment | 0.3% | 25.2% | 10.1% |
| (row %) | (60.5%) | (39.5%) | (100.0%) |
| Total | 1.1% | 44.2% | 17.2% |
| (row %) | (62.7%) | (37.3%) | (100.0%) |

Notes: This table presents the marriage rates by Round 3 enrollment status in Term 1, 2010 and treatment status. For each treatment arm, the top row summarizes the marriage rates by follow-up enrollment status, while the bottom row shows the sample size and indicates the raw follow-up row percentage in each cell. Means are weighted to make them representative of the target population in the study EAs.

Table IX: Teacher-Reported School Enrollment and Marital Status in Round 3

| | Dependent variable: | | | |
|---|---|---|---|---|
| | =1 if enrolled term 1 2010 | =1 if ever married | =1 if ever married | =1 if ever married |
| | All | All | Enrolled | Not Enrolled |
| | (1) | (2) | (3) | (4) |
| Conditional treatment | 0.058* | -0.026 | -0.012 | 0.033 |
| | (0.034) | (0.037) | (0.015) | (0.097) |
| Unconditional treatment | -0.000 | -0.088*** | -0.011 | -0.159** |
| | (0.036) | (0.030) | (0.010) | (0.067) |
| Mean in the control group | 0.598 | 0.199 | 0.017 | 0.469 |
| Sample Size | 844 | 844 | 490 | 354 |
| Prob > F(Conditional=Unconditional) | 0.099 | 0.106 | 0.857 | 0.088 |

Notes: The first two columns are impact regressions on Term 1, 2010 *teacher-reported* enrollment and being ever married by Round 3, respectively. Columns (3) and (4) present regressions of being ever married among those who were enrolled and not enrolled during Term 1, 2010, respectively. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table X: Heterogeneity of Program Impacts by Age-Group at Baseline

| | Dependent Variable: | | | |
|---|---|---|---|---|
| | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant |
| | (1) | (2) | (3) | (4) |
| Conditional treatment | 0.467*** | 0.141* | -0.023 | -0.008 |
| | (0.159) | (0.073) | (0.017) | (0.028) |
| Unconditional treatment | 0.257 | -0.116 | -0.051** | -0.059*** |
| | (0.157) | (0.102) | (0.020) | (0.020) |
| =1 if Over 15 | -0.786*** | -0.546*** | 0.122*** | 0.176*** |
| | (0.244) | (0.058) | (0.026) | (0.027) |
| Conditional treatment*Over 15 | 0.290 | 0.017 | 0.037 | 0.104* |
| | (0.291) | (0.089) | (0.056) | (0.054) |
| Unconditional treatment*Over 15 | 0.103 | 0.245** | -0.067 | -0.032 |
| | (0.255) | (0.110) | (0.042) | (0.046) |
| Number of unique observations | 852 | 2,057 | 2,084 | 2,087 |
| Prob > F(Conditional=Unconditional) | 0.095 | 0.031 | 0.188 | 0.067 |
| Prob > F(Conditional*Older=Unconditional*Older) | 0.364 | 0.059 | 0.097 | 0.027 |

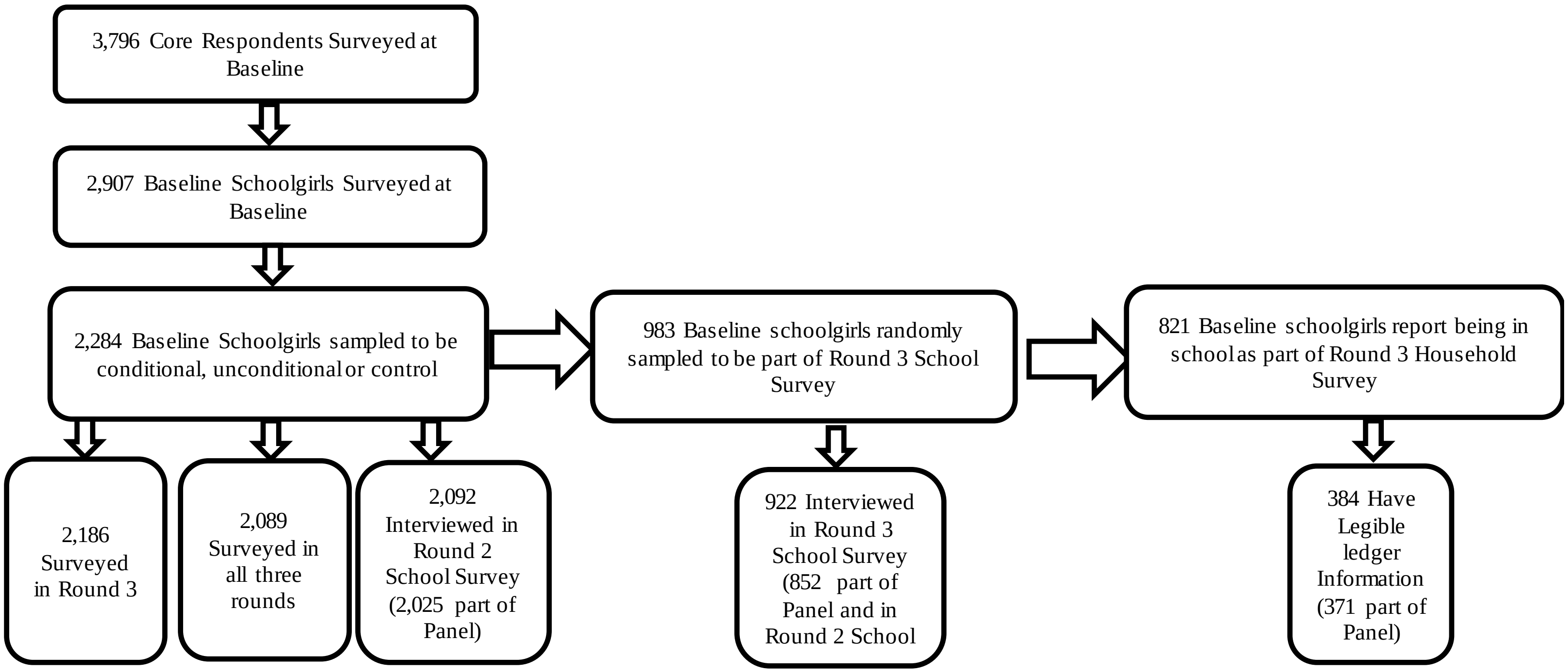
Notes: An indicator variable is constructed that takes on a value of one if the core respondent was older than fifteen at baseline, and is zero otherwise. Regressions are OLS models using Round 3 data with robust standard errors clustered at the EA level. Baseline values of the following variables are included as controls in the regression analyses: an indicator for being over fifteen, strata dummies, household asset index, highest grade attended, and an indicator for ever had sex. All regressions are weighted to make the results representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Table XI: Impacts of Household and Individual Transfer Amounts

| | Dependent variable: | | | |
|---|---|---|---|---|
| | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant |
| | (1) | (2) | (3) | (4) |
| Conditional treatment, individual amount | 0.024 | -0.032 | -0.002 | 0.006 |
| | (0.051) | (0.029) | (0.008) | (0.012) |
| Unconditional treatment, individual amount | -0.048 | -0.019 | -0.016 | 0.013 |
| | (0.064) | (0.038) | (0.011) | (0.013) |
| Conditional treatment, household amount | -0.027 | -0.000 | 0.001 | 0.005 |
| | (0.035) | (0.016) | (0.007) | (0.010) |
| Unconditional treatment, household amount | 0.081*** | -0.058** | -0.017** | -0.002 |
| | (0.031) | (0.029) | (0.007) | (0.009) |
| Conditional treatment, minimum transfer amounts | 0.572*** | 0.202* | -0.011 | 0.001 |
| | (0.213) | (0.118) | (0.044) | (0.052) |
| Unconditional treatment, minimum transfer amounts | 0.094 | 0.175 | 0.001 | -0.089* |
| | (0.167) | (0.132) | (0.040) | (0.050) |
| Number of unique observations | 852 | 2,057 | 2,084 | 2,087 |
| Prob > F(Conditional=Unconditional), individual amount | 0.390 | 0.788 | 0.300 | 0.702 |
| Prob > F(Conditional=Unconditional), household amount | 0.025 | 0.082 | 0.069 | 0.614 |
| Prob > F(Conditional=Unconditional), minimum amount | 0.046 | 0.877 | 0.834 | 0.203 |

Notes: Regressions are OLS models using Round 3 data with robust standard errors clustered at the EA level. 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 ever had sex. All regressions are weighted to make the results representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Figure I: Study Sample and Attrition



# Appendix A: 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.

Appendix Table B.1: Baseline Means and Balance for Round 3 School Survey Sample

| | Mean (s.d) | | | p-value (conditional-unconditional) |
|---|---|---|---|---|
| | Control group | Conditional group | Unconditional group | |
| **Panel A: Household-level variables** | (1) | (2) | (3) | (4) |
| Household size | 6.282 | 6.347 | 6.742** | 0.082* |
| | (2.167) | (2.120) | (2.129) | |
| Asset index | 0.456 | 0.771 | 1.313** | 0.271 |
| | (2.539) | (2.678) | (2.515) | |
| Female-headed household | 0.349 | 0.265* | 0.260* | 0.925 |
| | (0.477) | (0.442) | (0..440) | |
| Mobile phone access | 0.629 | 0.583 | 0.657 | 0.300 |
| | (0.484) | (0.494) | (0.476) | |
| Household transfer amount | N/A | 6.987 | 6.844 | 0.845 |
| | | (2.271) | (2.170) | |
| **Panel B: Individual-level variables** | | | | |
| Age | 15.288 | 15.005* | 15.324 | 0.044** |
| | (1.866) | (1.732) | (1.870) | |
| Highest grade attended | 7.425 | 7.272 | 7.831* | 0.014** |
| | (1.635) | (1.482) | (1.652) | |
| Mother alive | 0.853 | 0.804 | 0.835 | 0.368 |
| | (0.355) | (0.397) | (0.372) | |
| Father alive | 0.720 | 0.706 | 0.756 | 0.265 |
| | (0.450) | (0.456) | (0.430) | |
| Never had sex | 0.817 | 0.784 | 0.795 | 0.785 |
| | (0.387) | (0.412) | (0.405) | |
| Ever pregnant | 0.017 | 0.023 | 0.022 | 0.934 |
| | (0.131) | (0.150) | (0.147) | |
| Individual transfer amount | N/A | 2.970 | 2.936 | 0.725 |
| | | (1.438) | (1.415) | |
| Number of observations | 296 | 309 | 247 | |

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, while in column (4) stars indicate a significant difference between the conditional and unconditional treatment arms. Standard errors are clustered at the EA level to account for the design effect. Means are weighted to make them representative of the target population in the study EAs. The number of unique observations vary slightly across variables due to missing observations, with the total sample size ranging between 852 and 848.

Appendix Table B.2: Baseline Means and Balance for Round 3 Legible Ledger Sample

| | Mean (s.d) | | | p-value (conditional-unconditional) |
|---|---|---|---|---|
| | Control group | Conditional group | Unconditional group | |
| **Panel A: Household-level variables** | (1) | (2) | (3) | (4) |
| Household size | 6.122 | 5.982 | 6.684* | 0.022** |
| | (2.116) | (1.978) | (2.120) | |
| Asset index | 0.056 | 0.659 | 0.972** | 0.493 |
| | (2.459) | (2.658) | (2.206) | |
| Female-headed household | 0.318 | 0.307 | 0.220 | 0.265 |
| | (0.468) | (0.463) | (0.417) | |
| Mobile phone access | 0.604 | 0.547 | 0.443 | 0.403 |
| | (0.491) | (0.499) | (0.499) | |
| Household transfer amount | N/A | 6.966 | 6.788 | 0.763 |
| | | (2.113) | (1.819) | |
| **Panel B: Individual-level variables** | | | | |
| Age | 14.667 | 14.810 | 15.086 | 0.251 |
| | (1.516) | (1.639) | (1.861) | |
| Highest grade attended | 6.611 | 6.942 | 7.518*** | 0.069* |
| | (1.477) | (1.566) | (1.316) | |
| Mother alive | 0.896 | 0.782** | 0.855 | 0.135 |
| | (0.307) | (0.414) | (0.342) | |
| Father alive | 0.778 | 0.633* | 0.747 | 0.219 |
| | (0.417) | (0.484) | (0.437) | |
| Never had sex | 0.890 | 0.833 | 0.842 | 0.842 |
| | (0.314) | (0.374) | (0.367) | |
| Ever pregnant | 0.006 | 0.011 | 0.014 | 0.879 |
| | (0.080) | (0.106) | (0.116) | |
| Individual transfer amount | N/A | 2.940 | 3.188 | 0.346 |
| | | (1.407) | (1.508) | |
| Number of observations | 112 | 157 | 102 | |

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, while in column (4) stars indicate a significant difference between the conditional and unconditional treatment arms. Standard errors are clustered at the EA level to account for the design effect. Means are weighted to make them representative of the target population in the study EAs. The number of unique observations vary slightly across variables due to missing observations, with the total sample size ranging between 371 and 369.

**Appendix C: Robustness of enrollment findings with respect to school survey sampling and the treatment of missing values**

In Table III, we presented program impacts using self-reported (Panel A) and teacher-reported enrollment data (Panel B). The sample sizes for these analyses differ, and so do the samples sizes between 2008 and 2009 within Panel B. This appendix is provided to clarify the differences between the samples and to analyze the robustness of the findings in Table III to the choice of samples, as well as to the treatment of missing enrollment data from the school surveys.

Panel A in Table III uses the sample of 2,087 individuals that were interviewed in all three survey rounds (panel sample). Moving from Panel A to Panel B to report enrollment status reported by the teachers, there are two factors that cause the sample sizes to change. First, for some of the individuals who reported being enrolled in school, the enrollment data from the school survey are missing (we detail the reasons for why enrollment data are missing later in this appendix). Hence, the sample size for 2008 in Panel B (columns (1)-(3)) falls to 2,023 because teacher-reported enrollment data for 64 individuals is missing.[1] Second, we randomly sampled the individuals who were to be surveyed during the Round 3 school surveys.[2] Of the 983 individuals randomly selected from the original sample of baseline schoolgirls, 904 are also part of the panel sample. Of these 904 girls, teacher-reported enrollment data for either 2008 or 2009 are missing for 52 girls (23 in 2009 and 29 in 2008), which leaves 852 individuals for whom self-reported and teacher-reported enrollment data are available for both 2008 and 2009.[3] This is the sample that is used in columns (4)-(8) of Table III.[4]

Having explained that the sample sizes in Table III vary for two different reasons, we proceed to examine the robustness of our findings on enrollment with respect to each of them. First, we examine the least problematic reason for the differences in sample size; namely the random

[1] In column (4) of Table I, we have shown that the attrition rates for the Round 2 school survey are not significantly different between the control group and the two treatment arms, even though the likelihood of having a non-missing value for teacher-reported enrollment in the Round 2 school survey is approximately 6 pp higher in the CCT arm than the UCT arm (p-value=0.246).

[2] As also explained before in footnote 27 in the main text, the reason why the school survey in Round 3 was conducted for a randomly selected sub-sample is that school ledgers were also sought to check the attendance of core respondents. As locating these ledgers, examining them, and recording attendance for each core respondent is time consuming and costly, the study team decided to reduce the sample size.

[3] As our summary enrollment indicator (the total number of terms enrolled during 2008 and 2009) requires the availability of data from both years, we chose to restrict our sample in columns (4)-(8) to those for whom teacher-reported enrollment data are available for both years. The results are similar if we instead include those for whom 2009 data are available in columns (4)-(6) and column (8).

[4] In column (5) of Table I, we have shown that the attrition rates for the Round 3 school survey are not different between the control group and the two treatment arms. This remains true if we consider the 52 individuals for whom at least one of 2008 or 2009 enrollment data is missing.

sampling of those tracked during the Round 3 school surveys. Appendix Table C.1 reproduces Panel A from Table III and contrasts it with program impacts on self-reported enrollment using only the 904 individuals randomly selected for inclusion in the Round 3 school survey that are also part of the panel sample (Panel B). The standard errors in Panel B are substantially larger than those in Panel A due to the loss of power from the reduced sample size, so the UCT advantage in self-reported enrollment is no longer statistically significant in Panel B (p-value=0.241). Nonetheless the point estimates are broadly similar in the two tables, providing evidence that the sample that we attempted to track in the Round 3 school survey was representative.

We now move on to a careful analysis of the attrition that occurred in the sample originally selected to receive the Round 3 school survey, which concerns 52 individuals for whom we could not obtain teacher-reported schooling data. As explained in Section 2.4.1, using Round 2 (Round 3) household survey data, we collected the name of the school, the grade, and the teacher's name for the core respondent if she reported being enrolled in school at any point during the 2008 (2009) school year. These teachers were then located at the named schools and interviewed about each respondent's schooling status (term by term) during the past school year. The main reason for the attrition that occurred during this process was the fact that some of the schools in Zomba could not be located, either because they no longer were operating or, in a few cases, no one had heard of the school.[5] These constitute more than 70% of the cases with missing values for teacher-reported enrollment data.[6] Appendix Table C.2 presents the distribution of the reasons for missing school survey data by treatment status and shows that they differed somewhat between treatment arms: the girls in the CCT arm with missing school survey data were less likely to have reported being enrolled in a school that no longer existed or was unknown.

While we do not have enough information to definitively categorize any of these observations with missing teacher-reported enrollment data, we can test the robustness of our findings in Table III, by imposing several rules on how to treat these data if we do not want to leave them as missing values. We start with three simple rules that do NOT depend on treatment status:

[5] If the school no longer exists, we do not know the date on which it stopped operating, meaning that we cannot deduce whether the student who reported being enrolled at that school during the previous year is being truthful or not.

[6] In addition, some girls reported being enrolled in schools outside the study district of Zomba. We aggressively tracked these cases, but some of these schools could not provide us with information because the field teams did not have letters of introduction from the proper authorities for these schools, while, from others, the required information could not be obtained by phone. Finally, some schools in Zomba district refused to provide information while a few others were not visited due to (random) mistakes by our field office.

Rule 1. Set each missing value for teacher-reported enrollment equal to *zero* if the school no longer exists or is unknown, and to *one* otherwise,

Rule 2. Set all missing values for teacher-reported enrollment equal to *zeros (everyone lying),*

Rule 3. Set all missing values for teacher-reported enrollment equal to *ones (everyone truthful).*

Appendix Table C.3 presents program impacts on teacher-reported enrollment in 2008 and 2009 according to each of these rules.[7] We can see that the impact sizes in the CCT arm vary between 0.48 and 0.60 terms and are always significant at the 99 percent confidence level. In contrast, the impact sizes in the UCT arm vary between 0.20 and 0.27 and significant at the 90 percent confidence level only once – when we assume that all self-reports are truthful. The ratio of the impact in the UCT arm to the CCT arm is between 33% and 56% with the p-values for the differences between the two treatment arms equal to 0.03, 0.18, and 0.08, respectively. These findings confirm our conclusion from Table III that there was a large and significant improvement in school enrollment in the CCT arm and a substantially smaller improvement in the UCT arm that is barely significant at the 90 percent confidence level.

In addition to examining different rules for handling missing values, we can construct bounds for program impacts in each treatment arm by assigning opposite values of the dependent variable to observations in the two treatment arms and the control group (Lee, 2009). We present bounds using the following extreme cases (Control, CCT, and UCT):[8]

Rule 4. Control group lying, CCT truthful, and UCT lying (0, 1, 0),

Rule 5. Control group truthful, CCT lying, and UCT truthful (1, 0, 1),

Rule 6. Control group truthful, CCT truthful, and UCT lying (1, 1, 0),

Rule 7. Control group lying, CCT lying, and UCT truthful (0, 0, 1).

Appendix Table C.4 below shows that, as expected, these extreme assumptions over the nature of attrition cause fluctuation in the parameter estimates, but program impact on teacher-reported enrollment is significant in every case in the CCT arm, while the UCT effect varies between zero (Panel C) and being as large as the CCT effect (Panels B & D). In a total of eight runs (under which missing values were left missing or treated according to one of the seven rules

[7] Please note that the sample size is now 904, which includes all individuals who have been randomly sampled to be part of the Round 3 school survey and are part of the panel sample.

[8] With three groups (control, CCT, and UCT) and a binary enrollment variable (*zero* or *one*), there are eight combinations, two of which were covered above: Rule 2 (0, 0, 0) and Rule 3 (1, 1, 1). Of the remaining six combinations, we forego presenting two of them here, where the enrollment statuses assigned to the two treatment arms are concordant – (0, 1, 1) and (1, 0, 0) – as these are the least likely cases to affect the comparison of impacts in the CCT vs. the UCT arm.

imposed above), the program impacts in the CCT arm are always statistically significant – seven times at the 99 percent confidence level. In contrast, the impacts in the UCT arm were statistically insignificant in half of the runs, significant at the 90 percent confidence level twice, and significant above the 95 percent confidence level only when we assumed all those with missing data to be enrolled in school (the least likely scenario given that the UCTs appear most prone to over reporting their attendance). The advantage in the CCT arm is large and statistically significant in five of the eight runs, equally large with a p-value of 0.18 in another, and non-existent in the remaining two runs. Hence this bounding exercise confirms the fact that the school survey data reveal a sizable improvement in enrollment under the CCT and modest gains under the UCT.

In this appendix, we subjected the results presented in Table III to a battery of tests by varying the samples analyzed and the treatment of missing values. The analysis presented above supports the conclusions we reached using Table III: while self-reported data suggest that program impacts on enrollment were higher in the UCT than the CCT arm (or at the very least no worse), teacher-reported data suggest that the CCT arm outperformed the UCT arm – most likely by a large margin. Only under extreme assumptions that are contrary to other evidence available to us, we find that the improvements in enrollment in the two treatment arms were similar.

Appendix Table C.1: Program impacts on *self-reported* school enrollment

Panel A: Program impacts on *self-reported* school enrollment

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year 1: 2008 | | | Year 2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 Terms) | Term 1, Post-program |
| Conditional treatment | 0.007 | 0.019* | 0.041** | 0.049*** | 0.056*** | 0.061*** | 0.233*** | 0.005 |
| | (0.011) | (0.011) | (0.017) | (0.017) | (0.018) | (0.019) | (0.070) | (0.025) |
| Unconditional treatment | 0.034*** | 0.051*** | 0.054*** | 0.072*** | 0.095*** | 0.101*** | 0.406*** | 0.074*** |
| | (0.010) | (0.011) | (0.018) | (0.021) | (0.022) | (0.021) | (0.079) | (0.026) |
| Mean in Control | 0.958 | 0.934 | 0.900 | 0.831 | 0.800 | 0.769 | 5.191 | 0.641 |
| Number of observations | 2,087 | 2,087 | 2,086 | 2,087 | 2,087 | 2,087 | 2,086 | 2,086 |
| Prob > F(Conditional=Unconditional) | 0.006 | 0.012 | 0.460 | 0.299 | 0.102 | 0.098 | 0.038 | 0.028 |

Panel B: Program impacts on *self-reported* school enrollment using the Round 3 School Survey

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.017 | 0.018 | 0.055* | 0.032 | 0.057** | 0.076*** | 0.254** | 0.026 |
| | (0.018) | (0.018) | (0.031) | (0.022) | (0.028) | (0.025) | (0.104) | (0.033) |
| Unconditional treatment | 0.043** | 0.037** | 0.054* | 0.040 | 0.076** | 0.104*** | 0.353*** | 0.063* |
| | (0.017) | (0.018) | (0.030) | (0.026) | (0.030) | (0.031) | (0.115) | (0.033) |
| Mean in the control group | 0.947 | 0.947 | 0.900 | 0.865 | 0.817 | 0.777 | 5.253 | 0.652 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.055 | 0.137 | 0.954 | 0.714 | 0.450 | 0.273 | 0.241 | 0.278 |

Notes: The dependent variable in Panel A and Panel B is whether the student reported being enrolled in school for the relevant year/term. Panel A shows the results for the entire sample, while Panel B restricts the results to the sub-sample selected for the Round 3 School Survey. Post-program refers to Term 1, 2010, the first term after the program ended. Total terms refers to the total number of terms enrolled during the program. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Appendix Table C.2: Reason Why Teacher-Reported Enrollment Data is Missing by Treatment Status

| | Control | CCT | UCT | TOTAL |
|---|---|---|---|---|
| School no longer exists/unknown | 74.5% | 44.2% | 88.5% | 71.3% |
| N | 22 | 5 | 10 | 37 |
| School outside of Zomba district | 10.5% | 20.1% | 8.9% | 12.1% |
| N | 3 | 2 | 1 | 6 |
| Other (refused/coding error) | 15.0% | 35.7% | 2.7% | 16.6% |
| N | 5 | 4 | 0 | 9 |
| Total | 100.0% | 100.0% | 100.0% | 100.0% |
| N | 30 | 11 | 11 | 52 |

Notes: Means are weighted to make them representative of the target population in the study EAs.

Appendix Table C.3: Robustness of Teacher-Reported Enrollment to the Treatment of Missing Values

Panel A: Program impacts on teacher-*reported* school enrollment (Rule 1: zero if school does not exist, one otherwise)

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year 1: 2008 | | | Year 2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 Terms) | Term 1, Post-program |
| Conditional treatment | 0.084*** | 0.087*** | 0.104*** | 0.091** | 0.127*** | 0.108*** | 0.602*** | 0.078** |
| | (0.032) | (0.031) | (0.034) | (0.038) | (0.032) | (0.036) | (0.142) | (0.038) |
| Unconditional treatment | 0.012 | 0.031 | 0.013 | 0.035 | 0.066* | 0.043 | 0.199 | 0.039 |
| | (0.053) | (0.053) | (0.054) | (0.038) | (0.037) | (0.039) | (0.195) | (0.038) |
| Mean in Control | 0.850 | 0.826 | 0.803 | 0.757 | 0.729 | 0.702 | 4.666 | 0.578 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.155 | 0.280 | 0.081 | 0.133 | 0.037 | 0.062 | 0.031 | 0.232 |

Panel B: Program impacts on *teacher-reported* school enrollment (Rule 2: all missing values replaced with zeros)

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.072* | 0.074** | 0.092** | 0.086* | 0.121*** | 0.102*** | 0.547*** | 0.066* |
| | (0.038) | (0.037) | (0.040) | (0.046) | (0.036) | (0.038) | (0.175) | (0.040) |
| Unconditional treatment | 0.014 | 0.034 | 0.015 | 0.046 | 0.076** | 0.054 | 0.239 | 0.044 |
| | (0.054) | (0.055) | (0.054) | (0.038) | (0.037) | (0.039) | (0.199) | (0.037) |
| Mean in the control group | 0.841 | 0.817 | 0.794 | 0.744 | 0.716 | 0.689 | 4.602 | 0.569 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.318 | 0.498 | 0.195 | 0.392 | 0.191 | 0.205 | 0.175 | 0.522 |

Appendix Table C.3 cont.: Robustness of Teacher-Reported Enrollment to the Treatment of Missing Values

Panel C: Program impacts on teacher-reported school enrollment (Rule 3: all missing values replaced with ones)

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.049** | 0.052** | 0.069*** | 0.085** | 0.120*** | 0.101*** | 0.476*** | 0.078** |
| | (0.021) | (0.022) | (0.025) | (0.038) | (0.032) | (0.035) | (0.122) | (0.038) |
| Unconditional treatment | 0.037 | 0.056** | 0.038 | 0.032 | 0.063* | 0.040 | 0.266* | 0.039 |
| | (0.027) | (0.027) | (0.033) | (0.036) | (0.036) | (0.038) | (0.140) | (0.038) |
| Mean in the control group | 0.899 | 0.876 | 0.852 | 0.770 | 0.742 | 0.715 | 4.855 | 0.578 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.594 | 0.837 | 0.259 | 0.151 | 0.049 | 0.069 | 0.085 | 0.232 |

Notes: The dependent variable in all panels is whether the teacher reported the core respondent being enrolled in school for the relevant year/term. Post-program refers to Term 1 2010, the first term after the program ended. Total terms refers to the total number of terms enrolled during the program. All columns are restricted to the sub-sample of girls sampled for the Round 3 school survey who are also both part of the panel data set and part of the school survey panel. In Panel A, missing values are replaced with our "best guess" based on the reason for the missing data. In Panel B, all missing values are replaced with a value of one. In Panel C, all missing values are replaced with a zero. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Appendix Table C.4: Robustness of Teacher-Reported Enrollment to the Treatment of Missing Values (Lee Bounds)

Panel A: Program impacts on *teacher-reported* school enrollment (Rule 4: Control group lying, CCT truthful, and UCT lying (0, 1, 0))

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.113*** | 0.112*** | 0.129*** | 0.109*** | 0.144*** | 0.125*** | 0.739*** | 0.087** |
| | (0.029) | (0.028) | (0.030) | (0.038) | (0.032) | (0.036) | (0.133) | (0.038) |
| Unconditional treatment | 0.007 | 0.026 | 0.008 | 0.045 | 0.076** | 0.053 | 0.213 | 0.042 |
| | (0.054) | (0.055) | (0.054) | (0.038) | (0.037) | (0.039) | (0.196) | (0.038) |
| Mean in the control group | 0.841 | 0.817 | 0.794 | 0.744 | 0.716 | 0.689 | 4.602 | 0.569 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.033 | 0.089 | 0.016 | 0.085 | 0.018 | 0.035 | 0.004 | 0.176 |

Panel B: Program impacts on teacher-reported school enrollment (Rule 5: Control group truthful, CCT lying, and UCT truthful (1, 0, 1))

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.015 | 0.013 | 0.031 | 0.062 | 0.099*** | 0.083** | 0.295* | 0.056 |
| | (0.034) | (0.037) | (0.039) | (0.046) | (0.036) | (0.039) | (0.172) | (0.040) |
| Unconditional treatment | 0.051* | 0.069** | 0.050 | 0.033 | 0.065* | 0.045 | 0.303** | 0.041 |
| | (0.027) | (0.029) | (0.035) | (0.036) | (0.036) | (0.038) | (0.145) | (0.038) |
| Mean in the control group | 0.897 | 0.874 | 0.851 | 0.770 | 0.740 | 0.710 | 4.843 | 0.580 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.366 | 0.163 | 0.663 | 0.537 | 0.316 | 0.311 | 0.964 | 0.672 |

Appendix Table C.4 cont.: Robustness of Teacher-Reported Enrollment to the Treatment of Missing Values (Lee Bounds)

Panel C: Program impacts on teacher-reported school enrollment (Rule 6: Control group truthful, CCT truthful, and UCT lying (1, 1, 0))

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year1: 2008 | | | Year2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 terms) | Term 1, Post-program |
| Conditional treatment | 0.054*** | 0.050** | 0.066*** | 0.084** | 0.121*** | 0.106*** | 0.478*** | 0.076** |
| | (0.019) | (0.022) | (0.024) | (0.038) | (0.033) | (0.036) | (0.120) | (0.038) |
| Unconditional treatment | -0.047 | -0.029 | -0.047 | 0.020 | 0.052 | 0.032 | -0.031 | 0.034 |
| | (0.050) | (0.053) | (0.051) | (0.037) | (0.037) | (0.039) | (0.187) | (0.038) |
| Mean in the control group | 0.897 | 0.874 | 0.851 | 0.770 | 0.740 | 0.710 | 4.843 | 0.580 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.036 | 0.112 | 0.018 | 0.090 | 0.018 | 0.034 | 0.004 | 0.194 |

Panel D: Program impacts on teacher-reported school enrollment (Rule 7: Control group lying, CCT lying, and UCT truthful (0, 0, 1))

| | Dependent variable: =1 if enrolled in school during the relevant term | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Year 1: 2008 | | | Year 2: 2009 | | | Total Terms | Year 3: 2010 |
| | Term1 | Term2 | Term3 | Term1 | Term2 | Term3 | (6 Terms) | Term 1, Post-program |
| Conditional treatment | 0.074* | 0.075** | 0.094** | 0.087* | 0.122*** | 0.103*** | 0.555*** | 0.066* |
| | (0.038) | (0.037) | (0.040) | (0.046) | (0.036) | (0.038) | (0.175) | (0.040) |
| Unconditional treatment | 0.104*** | 0.124*** | 0.105*** | 0.058 | 0.089** | 0.066* | 0.547*** | 0.049 |
| | (0.033) | (0.032) | (0.039) | (0.037) | (0.037) | (0.038) | (0.152) | (0.038) |
| Mean in Control | 0.841 | 0.817 | 0.794 | 0.744 | 0.716 | 0.689 | 4.602 | 0.569 |
| Number of observations | 904 | 904 | 904 | 904 | 904 | 904 | 904 | 903 |
| Prob > F(Conditional=Unconditional) | 0.436 | 0.221 | 0.798 | 0.535 | 0.327 | 0.322 | 0.966 | 0.626 |

Notes: The dependent variable in all panels is whether the teacher reported the core respondent being enrolled in school for the relevant year/term. Post-program refers to Term 1 2010, the first term after the program ended. Total terms refers to the total number of terms enrolled during the program. All columns are restricted to the sub-sample of girls sampled for the Round 3 school survey who are also both part of the panel data set and part of the school survey panel. 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. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Appendix Table D.1: Heterogeneity of Program Impacts by Baseline Household Asset Index

| | Dependent Variable: | | | |
|---|---|---|---|---|
| | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant |
| | (1) | (2) | (3) | (4) |
| Conditional treatment | 0.585*** | 0.190*** | -0.010 | 0.030 |
| | (0.183) | (0.073) | (0.034) | (0.037) |
| Unconditional treatment | 0.392* | 0.046 | -0.086** | -0.054* |
| | (0.200) | (0.111) | (0.036) | (0.032) |
| =1 if Above Median Household Asset Index | 0.420 | 0.213*** | -0.033 | -0.031 |
| | (0.258) | (0.052) | (0.024) | (0.021) |
| Conditional treatment*Above Median Household Asset Index | -0.051 | -0.094 | -0.009 | -0.008 |
| | (0.317) | (0.095) | (0.035) | (0.044) |
| Unconditional treatment*Above Median Household Asset Index | -0.225 | -0.134 | 0.005 | -0.031 |
| | (0.313) | (0.110) | (0.048) | (0.048) |
| Number of unique observations | 852 | 2,057 | 2,084 | 2,087 |
| Prob > F(Conditional=Unconditional) | 0.306 | 0.248 | 0.083 | 0.059 |
| Prob > F(Conditional*Above Median Asset=Unconditional*Above Median Asset) | 0.493 | 0.761 | 0.775 | 0.696 |

Notes: An indicator variable is constructed that takes on a value of one if the core respondent's household asset index was above the median at baseline, and is zero otherwise. Regressions are OLS models using Round 3 data with robust standard errors clustered at the EA level. 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 ever had sex. All regressions are weighted to make the results representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

## Appendix E: Robustness Checks

*Spillover Effects of Assignment to Cash Transfer Locations*

In this experiment, treatment status with respect to conditionality was assigned at the EA level. Due to the proximity of EAs to each other, it is possible that the intermingling of students in the two treatment arms led to a change in behavior in the outcomes of interest, thus biasing our estimates of the marginal impact of the conditionality. One way of addressing this issue is to exploit the variation in treatment status across the locations at which the monthly cash transfers were made (Cash Transfer or CT locations). The CT location is the primary interface between beneficiaries and the program, so it provides a natural place to examine heterogeneity of program impacts. The locations were determined entirely by logistical concerns, and in many cases beneficiaries from multiple EAs were assigned to the same CT point. This variation can be informative because we would expect spillovers on the UCT arm to be stronger as the share of CCT beneficiaries at the CT point, for whom attendance is monitored and payments are withheld for non-compliance, increases.[9]

To test for this spillover of monitoring intensity, we calculate the number of girls assigned to each CT point (which is endogenous), the number of EAs that are serviced by each CT point (as a control for the heterogeneity of treatment status at a CT point, also endogenous), and the number of conditional girls assigned to each CT point (which is exogenous and random conditional upon the other two). The evidence from Appendix Table E.1 shows no evidence of any such spillovers: girls in the UCT arm do not behave differently when there is more intense monitoring of attendance around them. It appears unlikely that our estimates of differential program impacts are influenced by spillovers due to the proximity of girls with discordant treatment statuses.

[9] According to administrative program data, girls collected their own payments approximately 75% of the time. They also collected the payments for their parents/guardians roughly half of the time. These payments were made in separate envelopes to each designated recipient.

*Are the results robust to the handling of school fees under the program?*

As described in Section 2, secondary schools are not free in Malawi and the offers in the CCT arm included a promise to pay secondary school fees directly to the schools upon confirmation of enrollment by the program administrators.[10] To make the average transfer offers in the UCT arm equal to that in the CCT arm, the average school fee amount was added to the cash transfer offers of girls in the UCT arm who were eligible to attend secondary school at the beginning of the program (see footnote 22 for a more detailed description of this process). The relevant group eligible to attend secondary school at the outset of the program is those girls whose 'highest grade attended at baseline' are equal to Standard 8 or higher.

To test whether program impacts are influenced by the fact that school fee compensation was handled differently between the two treatment arms, we restrict our analysis to the sub-sample for whom school fees were not an issue: those whose highest grade attended at baseline was Standard 7 or lower. This group constitutes more than 56% of our study population. Columns (1)-(4) in Appendix Table E.2 show that all the impact findings are qualitatively the same as the average program impacts presented earlier, although there is less power due to the fact that sample size has been roughly halved. The CCT arm still holds an advantage in schooling outcomes, while the incidences of marriage and pregnancy are lower in the UCT arm.

Before the start of the second year of the program, all offers to those in the CCT and UCT arms were renewed. While those in the CCT arm who became eligible to attend secondary school still received offers for their secondary school fees to be paid, each offer in the UCT arm was identical to the previous one, meaning that girls who became eligible to attend secondary school were not offered additional payments in lieu of school fees.[11] This means that there is a group of

[10] Or, the students who paid their school fees could get reimbursed upon producing a receipt.
[11] While this could have been done in principle for everyone whose highest grade attended at baseline was Standard 7 (regardless of their actual school progress, which is endogenous to treatment), the study team opted not to do this for

girls in the UCT arm whose second year offers were smaller than their counterparts in the CCT arm. To examine whether this had an effect on differential program impacts, we rerun our impact regressions excluding this group from our sample, which constitutes approximately 25% of our target population. Columns (5)-(8) indicate that program impacts in this sub-sample are very similar to those for the entire sample.[12] We conclude that any influence of the way in which school fee compensation was handled in the two treatment arms on program impacts is likely to be very small.

fear of UCT girls 'sensing' that the additional payments were intended for school fees, and thereby contaminating the 'unconditional' treatment arm. As described in Section 2.3.2, during the first offers, UCT girls were told that they were randomly selected to receive bonuses as an explanation of their additional payments. The study team felt that this was not feasible in Year 2.

[12] Another way to address this question is to compare the two UCT cohorts that were just eligible to attend secondary school in Year 1 and Year 2. By comparing these two cohorts, we can examine whether girls in the UCT arm entering secondary school who did receive the additional payment fared better than the following cohort that did not. This analysis shows no sign of a differential impact for the cohort receiving the payment.

Appendix Table E.1: Spillover Effects of Assignment to Cash Transfer Locations

| | Dependent variable: | | | |
|---|---|---|---|---|
| | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant |
| | (1) | (2) | (3) | (4) |
| Number of Conditional Girls at CT Location | -0.005 | 0.001 | -0.000 | 0.001 |
| | (0.009) | (0.007) | (0.002) | (0.002) |
| Number of Baseline Schoolgirls at CT Location | -0.007* | -0.008*** | 0.001 | 0.001 |
| | (0.004) | (0.003) | (0.001) | (0.001) |
| Number of EAs served by CT Location | 0.020 | -0.002 | 0.008*** | 0.004 |
| | (0.013) | (0.016) | (0.003) | (0.003) |
| Number of unique observations | 247 | 258 | 261 | 261 |

Notes: The regressions are among the unconditional treatment group only. Regressions are OLS models with robust standard errors clustered at the EA level. All regressions are weighted to make the results 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 ever had sex. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.

Appendix Table E.2: Robustness of Results with Respect to Payments of Secondary School Fees

| | Core respondents in grade seven or below at baseline | | | | Excluding core respondents in grade seven at baseline | | | |
|---|---|---|---|---|---|---|---|---|
| | Dependent variable: | | | | | | | |
| | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant | Total number of terms enrolled (school survey) | Standardized English test score | =1 if ever married | =1 if ever pregnant |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Conditional treatment | 0.346** | 0.116* | -0.032 | 0.027 | 0.617*** | 0.127** | -0.000 | 0.035 |
| | (0.155) | (0.060) | (0.022) | (0.032) | (0.145) | (0.056) | (0.030) | (0.033) |
| Unconditional treatment | 0.183 | -0.048 | -0.058** | -0.075** | 0.285* | -0.035 | -0.088*** | -0.056* |
| | (0.182) | (0.118) | (0.029) | (0.031) | (0.159) | (0.075) | (0.025) | (0.032) |
| Number of unique observations | 524 | 1192 | 1,208 | 1,211 | 619 | 1,523 | 1,544 | 1,555 |
| Prob > F(Conditional=Unconditional) | 0.382 | 0.186 | 0.402 | 0.009 | 0.018 | 0.055 | 0.013 | 0.024 |

Notes: Columns (1)-(4) restrict analysis to the sample of core respondents in grade seven or below at baseline. Columns (5)-(8) exclude grade seven from the analysis. Regressions are OLS models using Round 3 data with robust standard errors clustered at the EA level. 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 ever had sex. All regressions are weighted to make the results representative of the target population in the study EAs. Parameter estimates statistically different than zero at 99% (***), 95% (**), and 90% (*) confidence.