Maira Emy Reimão

Cash and change

A replication study of a cash transfer experiment in Malawi

January 2019

Replication Paper 21

Multisector

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programs. We are the global leader in funding, producing and synthesizing high-quality evidence of what works, for whom, how, why and at what cost. We believe that using better and policy-relevant evidence helps to make development more effective and improve people's lives.

3ie Replication Paper Series

The 3ie Replication Paper Series is designed to be a publication and dissemination outlet for internal replication studies of development impact evaluations. Internal replication studies are those that reanalyze the data from an original paper in order to validate the results. The series seeks to publish replication studies with findings that reinforce or challenge the results of an original paper. To be eligible for submission, a replication study needs to be of a paper in 3ie's online Impact Evaluation Repository and needs to include a pure replication. 3ie invites formal replies from the original authors. These are published on the 3ie website together with the replication study.

The 3ie Replication Program also includes grant-making windows to fund replication studies of papers identified on our candidate studies list. Requests for proposals are issued one to two times a year. The candidate studies list includes published studies that are considered influential, innovative or counterintuitive. The list is periodically updated based on 3ie staff input and outside suggestions. The aim of the 3ie Replication Program is to improve the quality of evidence from development impact evaluations for use in policymaking and program design.

About this report

The Bill & Melinda Gates Foundation helped fund this report. All content, errors and omissions are the sole responsibility of the authors and do not represent the opinions of 3ie, its donors or its Board of Commissioners. Please direct any comments or queries to the corresponding author, Maira Emy Reimão, at maira.reimao@yale.edu.

Suggested citation: Reimão, ME, 2018. *Cash and change: a replication study of a cash transfer experiment in Malawi*, 3ie Replication Paper 21. Washington, DC: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/RPS0021

3ie Replication Paper Series executive editor: Marie Gaarder
Managing editor: Anna Heard
Production manager: Brigid Monaghan
Copy editor: Jaime L Jarvis
Proof reader: Yvette Charboneau
Cover design: John F McGill and Akarsh Gupta

Cash and change: a replication study of a cash transfer experiment in Malawi

Maira Emy Reimão
Yale University and Evidence Action

Replication Paper 21

January 2019

Summary

Unconditional cash transfers (UCTs) are increasingly held as a benchmark in development research and practice, against which, it is argued, other interventions should be compared for impact and cost-effectiveness. A 2011 study, *Cash or condition? Evidence from a cash transfer experiment*, by Baird and colleagues, is one of few studies to put this in practice and empirically compare, through a randomized controlled trial, the impact of a UCT to that of another treatment – in this case, a conditional cash transfer (CCT). This is a tremendous contribution to the literature on using UCTs as a comparison, as well as the literature on CCTs, and disentangling the roles of the transfer versus the conditionality.

A key contribution of this paper for the CCT literature lies in its analysis beyond outcomes directly related to the conditionality (school attendance) and into early marriage and pregnancies. The authors of the original study find that although the CCT increases school attendance (using teacher-reported data) and some test scores, the UCT decreases the likelihood of early marriage or pregnancy during the same two-year period, while the CCT does not. This puzzle is explained by the fact the conditionality of the CCT does encourage girls to stay in school, but for girls who drop out, the UCT continues to provide benefits, while the CCT does not. The transfer under the UCT arm to girls who drop out serves as a poverty alleviation tool that decreases the likelihood of other outcomes tied to low income, such as early marriage and pregnancy.

In this paper, we present a replication of *Cash or condition*. We are only able to access household survey data (and not information collected at the students' reported schools or results from tests administered by the original research team to participants) for the pure replication and measurement and estimation analysis. Given these data limitations, our study only replicates part of the paper. Nevertheless, our results are generally the same as in the original paper for the parts that are replicable, including the finding that the UCT decreases the likelihood of early marriage and pregnancy while the CCT does not. Although the original paper focuses on teacher-reported data, it also includes an analysis on student-reported enrollment and attendance. Limited to the latter set of data for this replication, we also find (as in the original paper but as opposed to using teacher-reported data) that the UCT increases self-reported enrollment and attendance to a greater extent than the CCT.

Apart from pure replication, we argue for addressing primary schoolgirls and secondary schoolgirls separately – or, alternatively, girls with stronger school attachment (i.e. closer to the expected grade for age) separately from girls with a history of repetition and/or dropouts (i.e. lower grade than expected for age). To study the potentially heterogeneous effects of UCT and CCT transfers for these sub-groups, we diverge from the original paper by running some analyses with sub-samples of the baseline schoolgirls. Here, we find that the UCT decreases the likelihood of pregnancy and marriage among primary schoolgirls and girls with a stronger attachment to school, but it decreases the likelihood of marriage only for secondary schoolgirls and those with a weaker school attachment. The impact on marriage for secondary schoolgirls is larger than for primary schoolgirls, although this is complicated by the fact that the effect is concentrated on girls who are the only beneficiaries in the household. In terms of policy, these results indicate that UCTs may be a powerful tool for decreasing pregnancy and

marriage rates among teenage girls, but particularly for decreasing the likelihood of pregnancy among girls in primary school and with a strong attachment to school. Our exploratory analysis also points to a warning regarding CCTs, which might increase pregnancies among older girls.

Contents

Summary .... **i**
List of tables .... **iv**
Abbreviations and acronyms .... **vi**
1. Introduction **1**
2. Replication process **3**
3. Push-button and pure replication results **4**
3.1 Push-button replication classification and justification 4
3.2 Pure replication output 4
3.3 Push-button and pure replication summary and wrap-up 12
4. Measurement and estimation analysis **13**
4.1 Measurement and estimation analysis results from proposed activities 14
4.2 Additional analysis 29
5. Limitations and other analysis originally in the research plan **39**
6. Conclusion **40**
Appendix: PBR output and comparison, by table **44**
References **54**

List of tables

Table 1: Pure replication of Table I 5
Table 2: Pure replication of Table II 6
Table 3: Pure replication of Table III, Panel A. 8
Table 4: Pure replication of Table VII 9
Table 5: Pure replication of Table VIII 10
Table 6: Pure replication of Table X 11
Table 7: Pure replication of Table XI 12
Table 8: Age distribution for baseline schoolgirls 14
Table 9: Age restricted to 13–18 – self-reported enrollment 15
Table 10: Age restricted to 13–18 – marriage and pregnancy 15
Table 11: Heterogeneous effects by age with cut-off at 18 16
Table 12: Age restricted to 13–18 – transfer amounts 16
Table 13: Grade for age 17
Table 14: Girls close to right grade for age – self-reported enrollment 18
Table 15: Girls close to right grade for age – marriage and pregnancy 18
Table 16: Girls below right grade for age – self-reported enrollment 19
Table 17: Girls below right grade for age – marriage and pregnancy 19
Table 18: UCT arm: number of girls by grade completed (reported at baseline) and additional transfers made 21
Table 19: Additional transfer dummy – self-reported enrollment 22
Table 20: Additional transfer dummy – marriage and pregnancy 22
Table 21: Accounting for additional transfer – transfer amounts 23
Table 22: Baseline primary schoolgirls – self-reported enrollment 24
Table 23: Baseline secondary schoolgirls – self-reported enrollment 24
Table 24: Baseline primary schoolgirls – marriage and pregnancy 25
Table 25: Baseline secondary schoolgirls – marriage and pregnancy 25
Table 26: Baseline primary schoolgirls – transfer amounts 26
Table 27: Baseline secondary schoolgirls – transfer amounts 26
Table 28: Additional effects on treated girls – self-reported enrollment 28
Table 29: Additional effects on treated girls – marriage and pregnancy 28
Table 30: Spillover effects on non-treated girls – self-reported enrollment 30
Table 31: Spillover effects on non-treated girls – marriage and pregnancy 30
Table 32: Self-reported school attendance – at least 4 out of 5 days 32
Table 33: Additional variables – baseline means and balance 32
Table 34: Additional variables – self-reported enrollment 34
Table 35: Additional variables – marriage and pregnancy 34
Table 36: Additional variables – age heterogeneity 35
Table 37: Additional variables – primary schoolgirls 36
Table 38: Additional variables – secondary schoolgirls 36
Table 39: Baseline primary schoolgirls – marriage and pregnancy 37
Table 40: Baseline secondary schoolgirls – marriage and pregnancy 38

Appendix tables

Table A1: PBR of Table I....44
Table A2: PBR of Table II....45
Table A3: PBR of Table III....46
Table A4: PBR of Table IV....47
Table A5: PBR of Table V....48
Table A6: PBR of Table VI....48
Table A7: PBR of Table VII....49
Table A8: PBR of Table VIII....50
Table A9: PBR of Table IX....51
Table A10: PBR of Table X....52
Table A11: PBR of Table XI....53

Abbreviations and acronyms

CCT	Conditional cash transfer
EA	Enumeration area
PBR	Push-button replication
UCT	Unconditional cash transfer

1. Introduction

Cash or condition? Evidence from a cash transfer experiment, by Baird and colleagues (2011), explores the impact of conditional and unconditional cash transfers on girls' school enrollment and attendance, learning, and likelihood of early marriage and pregnancies. The article is part of a larger research program, co-led by the authors of the paper, in the Zomba district of Malawi and based on offering cash transfers to girls aged 13 to 22 – conditional or unconditional on school attendance – over a two-year period (2008–2009).

Treatment assignment was done through a multi-step process. First, enumeration areas (EAs), each covering about 250 villages, were randomly assigned into an unconditional cash transfer (UCT) arm, a conditional cash transfer (CCT) arm or a control group. Then, within the UCT and CCT EAs, villages were randomly assigned an amount for the parent cash transfer. Finally, in the UCT and CCT EAs, each eligible girl was also randomly assigned an individual cash transfer amount. These randomly assigned monthly transfer amounts averaged US$7 for the parents and US$3 for the girls and, for the CCT arm, were conditional on the girls' school attendance (at least 80 percent) for the given month.

The key finding of *Cash and condition* is that CCTs increase school enrollment and attendance and UCTs do not, whereas UCTs lower girls' likelihood of getting pregnant or marrying early and CCTs do not. This puzzle is explained by the fact that CCTs encourage girls to comply with the condition of the transfer, but girls who do not comply and nonetheless drop out of school do not receive a transfer. In contrast, girls in the UCT arm who drop out of school continue to receive the transfer, relieving some of the financial pressure that might have otherwise encouraged them to marry and/or get pregnant early – or encouraged their parents to marry them off.

These results reveal that CCTs and UCTs may effect different behaviors and that the choice between the two depends on the desired outcome. CCTs may encourage compliance with the conditionality, but UCTs may have broader impacts through poverty alleviation. As UCTs are increasingly held as the benchmark against which other interventions are compared, this is a key contribution to the development literature and practice, as it shows that UCTs play a unique role, and their comparability to other programs depends on the target outcomes.

This is an important insight for development economics, in research and in policy, that speaks to several of the Sustainable Development Goals, including Goal 1 (no poverty), Goal 4 (quality of education) and Goal 5 (gender equality). In particular, it indicates that a UCT is an imperfect point of reference, since it may be a better tool for alleviating poverty and outcomes that flow from it, but not for changing and encouraging certain behaviors directly, such as girls' school enrollment and attendance.

For replicating *Cash or condition*, there are two publicly available data sets, each with different uses in the replication process. The first is a Stata do-file and clean data set that the original authors have made available through one of the authors' website.[1] This one is limited to the variables relevant to the original paper and, with the Stata do-file, designed for push-button replication (PBR). We refer to this as the "clean data set."

The second set of data can be found in the World Bank's Microdata Library. This one, archived as three separate data sets, covers three rounds of detailed household surveys, with several modules that were administered at the study baseline, midline and endline.[2] It does not, however, include information from surveys administered in schools (including teacher-reported attendance), scores from tests administered by the research team or information from school attendance ledgers, all three of which are also used in the analysis presented in *Cash or condition*. A code for transforming the three household data sets into the clean data set for PBR, or an identifier linking observations between the clean data set and this World Bank data set, does not appear to be publicly available.

Given its design and purpose, we use the clean data set for the PBR exercise. Although there are a few minor differences between the results presented in the original paper and those achieved through PBR, none of the discrepancies are meaningful, and most appear to be caused by a difference of a couple of observations between the two outputs.[3]

For pure replication and the measurement and estimation analysis, we use the data found in the World Bank's website, which includes all of the variables collected through the household surveys. To construct a data set for these two exercises, we combine the data across all three rounds. From this point on, we refer to this data set, using all three rounds of household survey data from the World Bank, as the "World Bank data set," and ignore all observations for girls who were not enrolled in school at baseline, as well as girls in treatment villages who were not directly treated ("spillover girls"[4]). These latter groups are included in other papers under this same research umbrella but excluded from the analysis for *Cash or condition*, which focuses on baseline schoolgirls offered conditional or unconditional offers. Reassuringly, the total number of girls in each of the three groups in the World Bank data set matches that described on page 8 of *Cash and condition*: 1,495 girls in control EAs; 506 girls assigned to the CCT arm; and 283 assigned to the UCT arm.

---

[1] Berk Özler's website (https://sites.google.com/site/decrgberkozler/papers-by-topic) includes replication files for many of his publications. The relevant files for *Cash or condition*, including the authors' memo, can be found at https://drive.google.com/open?id=0B274-JLBCKcdfnViOFhZbENJM1RYTGh4VkVwYThEajBCUUMzT2l3eHM2RFNPUWxKWUg4dnc.

[2] The data for each of the three rounds can be found at http://microdata.worldbank.org/index.php/catalog.

[3] In the public replication folder, along with the two replication files for *Cash or condition* (do-file and clean data set), the authors included a memo discussing a few discrepancies between the results presented in the paper and the output they obtain when running the do-file they share for replication. The differences between the results in their paper and their memo are small and do not change any of the interpretations, but an explanation for these discrepancies is not given in the memo. They do note, however, that the results in the memo are based on a "clean data set." As *Cash or condition* is one of many papers relating to their larger project, it is possible that the data set was further cleaned after the publication of the results.

[4] We re-visit this group briefly in Section 4.2, when exploring spillover effects.

2. Replication process

For PBR, we use the files available through Özler's website. Once the paths were changed, the do-file ran smoothly and produced all of the expected tables in the original paper. The one exception is Table IV, which, as the authors remark in the original paper, is based on a table from a different paper, written by a subset of the co-authors (Baird and Özler 2012). For this table, we use the clean data set and do-file from the latter paper, also available on Özler's website. The tables obtained through this PBR exercise are then compared to those presented in the published version of *Cash or condition*, as discussed in the next section.

To build the World Bank data set, we merge across various questionnaire modules in each round and then append the data over rounds, creating a data set with similar structure to the one provided by the authors for PBR, with three observations per individual (one for each round). While most of the variables from this data set can be used directly for the replication exercise, we construct a few of them, such as the asset index and highest school grade attended. We use the first principal components score to build the former,[5] while the latter relies on two separate questions in the survey (highest schooling level attended and years spent at that level).[6] The age of the study girls is also derived from two questions in the household roster. In particular, the month and year of birth was recorded for each household member, or, alternatively, their age. We use the recorded age, if provided (and within the allowed range of 13–22 years), and otherwise use the former to calculate their age at the time of the baseline survey.[7]

Crucially, however, the World Bank data set is based only on the household survey data (which is all that is provided in the World Bank Microdata Library), and therefore does not have all of the variables included in the clean data set. In particular, the World Bank data set does not include information from the surveys administered at the girls' reported schools in rounds 2 and 3, the scores from tests administered by the research team or attendance records found through school ledgers. We were also unable to obtain these from the original authors after a few requests.[8] As such, some of parts of the planned replication cannot be carried out, as discussed in Section 5.

[5] More specifically, we use indicators of ownership for each of the 25 durable goods included in Section 3 of the household baseline questionnaire and replace any missing values with the mean for that item before extracting the score. Imputing the mean of a variable when its value is missing is a standard approach for principal component analysis but has the drawback of reducing the variance of those variables, with potential repercussions for the analysis. This weakness is less problematic if there are few missing values, as is the case here (Little and Rubin 2002). We do not explore alternate specifications for constructing the asset index but note that not imputing missing asset values would yield 18 missing values for the asset index across all study girls.

[6] These two questions appear in the household and individual questionnaires at baseline. To construct this variable, we first use the data from the individual questionnaire. Then, for girls for whom the data is missing, we complete it with the information provided in the household questionnaire. Ten missing observations are filled in through the household questionnaire data.

[7] Twenty-eight girls out of 3,790 have a missing entry for age once these two steps are employed.

[8] The clean data set does not include respondent ID codes in a way that allows us to merge it with the World Bank data set to get these additional variables either. The replication team contacted

In some cases, where we state that we were unable to replicate certain tables or columns because they rely on teacher-reported attendance rates (e.g. Tables VIII, IX and parts of X and XI), it may have been possible to use self-reported information to produce the respective output instead. We have chosen not to do so in the pure replication, however, as a principal argument in the original paper – and in a companion piece by a subset of the research team (Baird and Özler 2012) – is that self-reported data is less reliable than teacher-reported data for the purpose of this study.

We use the World Bank data set for the measurement and estimation analysis as well, largely following the initial proposal for this replication (Reimão 2017). The absence of teacher-reported information and test scores from this data set, however, limits the scope of this activity to analyses that use self-reported enrollment and/or marriage and pregnancy as outcomes.

3. Push-button and pure replication results

3.1 Push-button replication classification and justification

This PBR is classified as complete with minor differences, with the following justification: with small exceptions, all of the results presented in the original paper are replicable with the publicly available do-files and cleaned data sets for the paper. The exceptions include a portion of Table II, which is nonetheless easily calculable, and Table IV, a version of which can be replicated with a different do-file and data set.

In general, there are a few differences between the results presented in the original paper and that obtained through the PBR exercise, and most of them are likely due to slight differences in the number of observations used. Also, even though significance levels differ (with conventional cutoffs) for a handful of coefficients between the two outputs, none of the results are meaningfully different. There are no substantial or qualitatively different findings between those published in the original paper and those achieved through the PBR exercise.

The tables produced through the PBR exercise, along with a comparison to the original tables in the paper, are included as an appendix.

3.2 Pure replication output

In this section, we summarize our pure replication results for each table in the original paper. As a reminder, the pure replication is done using the World Bank data set, not the clean data set, as the former is closest to the “raw” data. We present all of the respective tables we were able to produce through pure replication. When our output is not identical to that in the original paper, the original results are in black and the pure replication output is **bolded**.

the authors of *Cash or condition* to request the data from school surveys (teacher-reported), test scores and school attendance ledgers that are used in the paper but not publicly available outside the PBR data set, but the team did not receive the data by the time this paper was completed.

3.2.1 Table I – partial

The first table in *Cash or condition* addresses attrition over the three rounds of the study (baseline, midline and endline) and shows that any losses in observations are not correlated with treatment arm. We are able to replicate only the first two columns of Table I and achieve results identical to those in the original paper, confirming no differences by treatment in the likelihood of participating in the third household survey round (endline) or of being part of the complete panel data. We are not able to replicate the subsequent four columns of the table, as the outcome variables used for them are not in the World Bank data set. These variables are: whether the respondent took the tests administered by the research team; their information was collected through the school surveys (rounds 2 and 3); and they had legible school ledgers.

Table 1: Pure replication of Table I

Analysis of attrition		
	(1) Surveyed in round 3	(2) Surveyed in all three rounds
= 1 if conditional schoolgirl	0.020	0.021
	(0.015)	(0.030)
= 1 if unconditional schoolgirl	0.021	0.030
	(0.019)	(0.024)
Control group mean	0.946	0.893
	(0.009)	(0.011)
Number of observations	2,284	2,284
Prob > F (conditional = unconditional)	0.965	0.797

Note: Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.2.2 Table II – complete

Table II shows the balance across arms (control, CCT and UCT) in terms of both household- and individual-level variables. We are able to replicate all of Table II, and our results are generally consistent, though not identical, to those presented in the original paper.

Two small differences can be found in the descriptive statistics for the asset index (which we constructed ourselves in the replication; see footnote 5) and mean age. In particular, the asset index for the UCT arm in our analysis appears to be different from the control group at the 90 percent confidence level, although still not different from the CCT arm; and the difference in mean age in the treatment arms, compared to the control, are more significant in our results than in the original paper, besides being different across the two treatment arms.

Nevertheless, these differences do not lead to meaningful issues later in the analysis, as the asset index and the age of each girl at baseline are both included as control variables in the regressions carried out throughout the paper. The difference in the share of female-headed households between the control and treatment arms is also stronger in our replication compared to the results presented in the original paper, although there is no significant difference between the two treatment arms.

An additional remark regarding the control variables to be made here is that the variable referred to as “highest grade attended” in the original paper (and in this replication) is not strictly so. Rather, in the questionnaire, respondents were asked about the highest level of schooling attended (e.g. primary or secondary) as well as the years spent in that level. Considering that this is a context with high levels of repetition, this variable is likely upwardly biased, although with no reason to expect any difference in this bias between treated and non-treated arms at baseline.

Table 2: Pure replication of Table II

	Baseline means and balance			
	(1) Control group mean	(2) = 1 if conditional schoolgirl	(3) = 1 if unconditional schoolgirl	(4) p-value (conditional-unconditional)
Panel A: Household-level variables				
Household size	6.432 **6.437**	6.384 **6.330**	6.662 **6.604**	0.202 0.179
	(2.257) **(2.295)**	(2.146) **(2.136)**	(2.075) **(2.051)**	
Asset index	0.581 **0.523**	0.984 **0.923**	1.221 **1.165***	0.623 **0.603**
	(2.562) **(2.457)**	(2.740) **(2.621)**	(2.447) **(2.399)**	
Female-headed household	0.343 **0.351**	0.252** **0.256*****	0.245** **0.241*****	0.899 **0.740**
	(0.475) **(0.477)**	(0.434) **(0.437)**	(0.431) **(0.428)**	
Access to mobile phone	0.616 **0.623**	0.583 **0.599**	0.605 **0.626**	0.799 **0.752**
	(0.487) **(0.485)**	(0.494) **(0.491)**	(0.490) **(0.485)**	
Transfer amount to household (kwacha)	n/a	6.991 **7.045**	6.829 **6.906**	0.822 **0.846**
		(2.319)	(2.101) **(2.107)**	
Panel B: Individual-level variables				
Age	15.252 **15.305**	14.952* **14.990****	15.424 **15.519***	0.007*** **0.001*****
	(1.903) **(1.958)**	(1.827) **(1.836)**	(1.923) **(1.927)**	
Highest grade attended	7.478 **7.521**	7.246 **7.252**	7.896** **7.950****	0.004*** **0.001*****
	(1.634) **(1.669)**	(1.598) **(1.630)**	(1.604) **(1.609)**	
Mother still alive	0.842 **0.834**	0.802 **0.792**	0.836 **0.833**	0.360 **0.248**
	(0.365) **(0.372)**	(0.399) **(0.406)**	(0.371) **(0.373)**	
Father still alive	0.705 **0.698**	0.714 **0.716**	0.759 **0.755**	0.288 **0.355**

Baseline means and balance				
	(1) Control group mean	(2) = 1 if conditional schoolgirl	(3) = 1 if unconditional schoolgirl	(4) p-value (conditional-unconditional)
	(0.456) **(0.459)**	(0.453) **(0.451)**	(0.428) **(0.431)**	
Never had sex	0.797 **0.794**	0.797 **0.801**	0.775 **0.791**	0.582 **0.792**
	(0.402) **(0.405)**	(0.403) **(0.399)**	(0.419) **(0.407)**	
Ever pregnant	0.023 **0.021**	0.030 **0.028**	0.031 **0.028**	0.973 **0.992**
	(0.149) **(0.142)**	(0.171) **(0.166)**	(0.173) **(0.166)**	
Transfer amount to treated girl (kwacha)	n/a	3.090 **3.107**	3.033 **3.051**	0.606 **0.579**
		(1.431) **(1.415)**	(1.451) **(1.459)**	
Observations	1356 **1483-1491**	470 **500-506**	261 **269-283**	

3.2.3 Table III – partial

The third table is a key element of the paper, as it shows two sets of results. First, according to self-reported information, both CCTs and UCTs have a positive impact on school enrollment during the treatment period. Second, however, according to teacher-reported enrollment, only CCTs tend to have a positive impact on school enrollment during the same period.

Unfortunately, with the World Bank data, we are able to replicate Panel A in Table III, which uses student-reported enrollment, but not Panel B, as it is based on teacher-reported school enrollment. Our results for Panel A are generally the same as in the original paper, in terms of the magnitude and significance levels of the coefficients, but are not identical, as we have about 10 fewer observations for each column.

However, the question in the survey from which the dependent variable is extracted is not set in terms of enrollment per se but as, “were you attending school any time during Term [x]?” In other words, in our replication, we find that UCT and CCT girls are both more likely to attend school at any point during the treatment period, or to at least report doing so. There is no significant difference in self-reported attendance between the two treatment arms.

Without the data for reproducing Panel B, we cannot verify that this pattern does not hold in the teacher-reported data, with only CCT girls actually reported by teachers as enrolled at higher levels than the control group during the treatment period.

Table 3: Pure replication of Table III, Panel A

	Enrollment – self-reported							
	(1) 2008 Term 1	**(2) 2008 Term 2**	**(3) 2008 Term 3**	**(4) 2009 Term 1**	**(5) 2009 Term 2**	**(6) 2009 Term 3**	**(7) TOTAL number of terms**	**(8) 2010 Term 1**
= 1 if conditional schoolgirl	0.007 **0.006**	0.019* **0.018***	0.041** **0.040****	0.049***	0.056*** **0.054*****	0.061*** **0.059*****	0.233*** **0.227*****	0.005 **0.003**
	(0.011)	(0.011)	(0.017)	(0.017)	(0.018) **(0.019)**	(0.019)	(0.070)	(0.025) **(0.026)**
= 1 if unconditional schoolgirl	0.034*** **0.032*****	0.051*** **0.047*****	0.054*** **0.050*****	0.072***	0.095*** **0.099*****	0.101*** **0.104*****	0.406***	0.074***
	(0.010)	(0.011) **(0.010)**	(0.018) **(0.017)**	(0.021) **(0.022)**	(0.022) **(0.023)**	(0.021)	(0.079)	(0.026)
Mean of control group	0.958	0.934	0.900	0.831	0.800	0.769	5.191	0.641
Number of observations	2,087 **2,076**	2,087 **2,076**	2,086 **2,075**	2,087 **2,076**	2,087 **2,076**	2,087 **2,076**	2,086 **2,076**	2,086 **2,076**
Prob > F (conditional = unconditional)	0.006	0.012 **0.021**	0.460 **0.532**	0.299 **0.313**	0.102 **0.060**	0.098 **0.058**	0.038 **0.030**	0.028 **0.022**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.2.4 Tables IV–VI – could not be replicated

In *Cash or condition*, Table IV comes from a different paper by a subset of the same authors (Baird and Özler 2012), which argues that the teacher-reported data on school attendance is more reliable than self-reported data. To show this, the authors compare the two sources of data to administrative information recorded in school ledgers. Assuming the accuracy of the latter, they show that girls in the UCT and control groups are more likely to over-report attendance (relative to the ledger) than those in the CCT group. In contrast, there does not appear to be a systematic difference in teacher over-reporting across arms.

Table V then shows results similar to those in Panel B of Table III, indicating that the CCT arm increases the fraction of days attended, based on school ledger data, whereas the UCT does not. Table VI corroborates the higher level of attendance by showing higher test scores in English comprehension and cognitive ability for the girls in the CCT arm but not those in the UCT arm, relative to those in the control group.

Using the World Bank data set does not allow us to replicate any of these three tables, as Tables IV and V use data from school ledgers, whereas Table VI uses scores from tests administered by the researchers.

3.2.5 Table VII – complete

The central finding presented in *Cash or condition* is as follows: although the CCT has a stronger impact on compliance with the conditionality itself (i.e. school attendance), the

UCT has an effect on other desirable outcomes. Indeed, Table VII in the original paper shows that the UCT decreased marriage and pregnancy rates by round 3, but the CCT did not. We are able to replicate this table with similar, though not identical, results. The magnitudes and statistical significance of the coefficients achieved through our analysis are essentially the same as those in the original paper, showing negative impacts on marriage and pregnancy rates for the UCT arm but not the CCT arm, especially at the end of the two-year treatment (round 3).

Table 4: Pure replication of Table VII

	Marriage and pregnancy			
	(1) Ever married R2	**(2) Ever married R3**	**(3) Ever pregnant R2**	**(4) Ever pregnant R3**
= 1 if conditional schoolgirl	0.007 **0.008**	–0.012 **–0.002**	0.013 **0.014**	0.029 **0.041**
	(0.012) **(0.013)**	(0.024)	(0.014)	(0.027)
= 1 if unconditional schoolgirl	–0.026**	–0.079*** **–0.076*** **	–0.009 **–0.006**	–0.067*** **–0.069*** **
	(0.012) **(0.011)**	(0.022) **(0.021)**	(0.017)	(0.024) (0.023)
Mean of control group	0.043	0.180 **0.176**	0.089	0.247
Number of observations	2,087 **2,076**	2,084 **2,076**	2,086 **2,075**	2,087 **2,076**
Prob > F (conditional = unconditional)	0.024 **0.022**	0.025 **0.010**	0.265 **0.329**	0.003 **0.001**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.2.6 Table VIII – partial

To explain the positive impact of the UCT (but not the CCT) on marriage and pregnancy rates, the authors of the original study present descriptive statistics on the marriage rates of girls in each treatment arm, split by whether they are still enrolled in school by 2010 (the end of the treatment). Table VIII shows that girls in the UCT arm who have dropped out of school have lower marriage rates than girls in the control group or the CCT arm who are also no longer enrolled in school.

Unfortunately, Table VIII uses teacher-reported enrollment to sort girls enrolled and no longer enrolled in school by 2010. As a result, we cannot fully replicate it with the World Bank data set.

We can only replicate column 3 in this table (the aggregate column), although our results show higher marriage rates for each of the groups. Nevertheless, they are consistent with the findings in the respective column in the original paper – that girls in the UCT arm have lower marriage rates in round 3 than do girls in the other two arms. That said, we do not have the data to replicate the principal argument of this table, which is that this difference is primarily driven by lower marriage rates among dropout girls in the UCT arm.

Table 5: Pure replication of Table VIII

Marital status by enrollment and group			
		(3) **Total**	
Control, %	19.9		**21.4**
(row %)		(100.0)	
Conditional treatment, %	16.0		**21.5**
(row %)		(100.0)	
Unconditional treatment, %	10.1		**13.1**
(row %)		(100.0)	
Total, %	17.2		**20.4**
(row %)		(100.0)	

3.2.7 Table IX – could not be replicated

The first two columns in Table IX reiterate findings presented above – that UCT girls are not more likely to be enrolled in school by 2010 according to teacher-reported data (as in Table III, panel B), but are less likely to be married by round 3 (as in Table VII). The next two columns conduct the same regressions for girls who are enrolled and not enrolled in school (teacher-reported), and show that within the not-enrolled group, girls in the UCT arm are less likely to be married, whereas there is no difference in marriage rates within the enrolled group.

This table cannot be replicated using the World Bank data set, as it relies on teacher-reported enrollment data.

3.2.8 Table X – partial replication

Next, the authors look at heterogeneous effects by age. In Table X, they present results for similar regressions discussed previously, but with a dummy variable for girls older than 15 years and an interaction term between this dummy and the two treatment arms. The first two columns of Table X cannot be replicated, as they use teacher-reported enrollment and English comprehension scores as the outcome variables.

We replicate the other two columns in Table X. Although not identical, the magnitudes and statistical significance of the coefficients are similar to those presented in the original paper. In particular, we find slightly stronger (more significant) differences between the coefficients in the UCT and CCT arms and weaker differences between the coefficients on these treatments interacted with age.

Based on these results, it would appear that the UCT treatment decreases the likelihood of marriage and of pregnancy and, not surprisingly, that older girls are more likely to be married or get pregnant (although we find this latter difference to be smaller in magnitude than in the original paper). A differentiated impact on the likelihood of ever being pregnant among older girls in the CCT arm is detected only at the 10 percent level.

Table 6: Pure replication of Table X

Age heterogeneity				
	(3) Ever married		(4) Ever pregnant	
= 1 if conditional schoolgirl	–0.023	**–0.016**	–0.008	**0.007**
	(0.017)	**(0.018)**	(0.028)	
= 1 if unconditional schoolgirl	–0.051**	**–0.052*** **	–0.059***	**–0.063*** **
	(0.020)		(0.020)	**(0.020)**
Conditional treatment *above 15 years old	0.037	**0.040**	0.104*	**0.090***
	(0.056)		(0.054)	**(0.053)**
Unconditional treatment * above 15 years old	–0.067	**–0.052**	–0.032	**–0.019**
	(0.042)	**(0.037)**	(0.046)	**(0.047)**
= 1 if above 15 years old	0.122***	**0.097*** **	0.176***	**0.121*** **
	(0.026)	**(0.025)**	(0.027)	
Number of observations	2,084	**2,076**	2,087	**2,076**
Prob > F (conditional = unconditional)	0.188	**0.107**	0.067	**0.016**
Prob > F (conditional * older = unconditional * older)	0.097	**0.133**	0.027	**0.073**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.2.9 Table XI – partial replication

In the final table, the authors consider heterogeneous effects based on the value of the transfers offered to the household and the girls. Because they use the same outcome variables as those in Table X, we can again only replicate the final two columns in Table XI.

Nevertheless, the two columns we can replicate yield results that are qualitatively the same as those presented in the original paper, with the household amount given under the unconditional treatment as the only treatment variable with an impact on the likelihood of marriage. In contrast, the only coefficient picked up as significant (at conventional levels) for the likelihood of pregnancy is selection into the UCT arm.

Table 7: Pure replication of Table XI

Transfer amounts				
	(3)		**(4)**	
	Ever married		**Ever pregnant**	
Conditional treatment, individual amount	–0.002	**–0.001**	0.006	**0.001**
	(0.008)		(0.012)	**(0.015)**
Unconditional treatment, individual amount	–0.016	**–0.017**	0.013	**0.012**
	(0.011)		0.013	**(0.014)**
Conditional treatment, household amount	0.001	**0.000**	0.005	**0.006**
	(0.007)		(0.010)	
Unconditional treatment, household amount	–0.017**	**–0.015****	–0.002	
	(0.007)	**(0.006)**	(0.009)	
Conditional treatment, minimum transfer	–0.011	**–0.000**	0.001	**0.020**
	(0.044)	**(0.043)**	(0.052)	**(0.055)**
Unconditional treatment, minimum transfer	0.001	**0.002**	–0.089*	**–0.089***
	(0.040)	**(0.040)**	(0.050)	**(0.052)**
Number of observations	2,084	**2,076**	2,087	**2,076**
Prob > F (conditional = unconditional), individual	0.300	**0.250**	0.702	**0.621**
Prob > F (conditional = unconditional), household	0.069	**0.099**	0.614	**0.589**
Prob > F (conditional = unconditional), minimum	0.834	**0.970**	0.203	**0.147**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.3 Push-button and pure replication summary and wrap-up

Given the data sets that are publicly available, we can only partially replicate the *Cash or condition* paper in a pure replication exercise, although all of it is replicable under a PBR. The portions that are replicable produce results that are largely the same as those presented in the original paper. Although our sample in the pure replication often has about 10 observations fewer than those in the paper, the magnitudes and statistical significance of the coefficients are very similar to those in the original tables. A few exceptions include those in the balance tests (Table II), but most of the variables in question are included as controls in the regressions, so these small discrepancies do not meaningfully change subsequent results.

That said, the data sets that are not publicly available are crucial for the replication exercise, in the sense of verifying the main findings and message of the original paper. Limited to the World Bank data set, our replication only supports the second half of the principal result in *Cash or condition*: that the UCT treatment decreases marriage rates by 8 percentage points and the likelihood of ever having been pregnant by 7 percentage points by the end of the two years. Without the teacher-reported enrollment (or school ledger) data, we cannot confirm that this difference is driven by girls who drop out of school, much less see that the unique advantage of the CCT is in raising attendance rates.

In fact, with the data we have and the parts we have been able to replicate in this exercise, it would appear that CCTs and UCTs both raise school attendance rates (self-reported; Table III, part A), with the UCT leading to larger effects in some cases. The benefits of the UCT in delaying marriage and pregnancies are in addition to this impact. As a whole, then, the conclusion from the pure replication is drastically different without the teacher-reported data. When limited to self-reported enrollment data, the UCT appears to be a more impactful intervention – not only for increasing attendance rates, but also for delaying marriage and pregnancy – than the CCT.

4. Measurement and estimation analysis

In this section, we present the measurement and estimation analysis, as defined in Brown and colleagues (2014) for the use of existing data, including robustness checks and an exploration of heterogeneous effects conducted, given the target group for the study and the program design. Following the original replication plan (Reimão 2017), in Section 4.1 we look into the role of age in the recorded effects. We first focus on girls in the sample who are still of school age, and then split the sample by whether a girl is close to the right "age for grade" – the grade level expected for her age, assuming normal progress. We also explore the importance of a detail in the design of the program, whereby girls in the CCT arm who were in secondary school at baseline had their school fees covered, while the parents of those in the UCT arm received an additional transfer that could be used to cover school fees but was not explicitly presented to recipients as such.

In this section, our analysis tends to focus on marriage and pregnancy as the main outcomes, because these are the ones that, along with self-reported school attendance, are available in the World Bank data set and are studied in the original paper. We also conduct some analysis using self-reported school attendance and include those results here, although we point again to the original authors' argument these data are less reliable than teacher-reported results.

In addition to that proposed in the replication plan, we conduct some exploratory analysis as we became more familiar with the study's survey instruments and design.[9] For example, as we learned of the variables collected at baseline, we considered the role of certain variables that appeared to be potentially relevant to the main outcomes in this replication – marriage and pregnancy (e.g. a girl's reported desire to ever be married or having a close friend who has had sex). Results for all of this supplementary work are presented in Section 4.2, separate from those that were originally proposed, and should be considered exploratory.

[9] The replication plan was written after we had downloaded the clean data set and the World Bank data set but had not run any analysis, including gathering descriptive statistics or thoroughly reading through the survey instruments.

As with the pure replication, this measurement and estimation analysis is constrained by the available data. Some activities proposed in the replication plan could not be carried out, as they require teacher-reported school attendance data and/or student test scores. We address these limitations in Section 5.

4.1 Measurement and estimation analysis results from proposed activities

4.1.1 Age and age-for-grade

As described in the original paper, the study includes girls aged 13–22 at baseline. A concern about such a broad age range is that the older girls in this bracket are no longer school-aged, and may thus behave differently to a cash transfer than school-aged girls. In particular, we might expect that older girls are less likely to attend school (and/or more likely to drop out) and perhaps also less likely to respond to a cash transfer by attending school. Alternatively, even though the individual portion of the cash transfer is given directly to girls, younger girls may have less say in their school attendance (or marriage) than girls who are past their teenage years. This distinction may also be relevant for policy, as it is likely that a program designed to encourage school enrollment and attendance, when faced with limited resources, would target school-aged girls rather than those 19 and older.

It turns out, however, that girls older than 18 make up less than 10 percent of all girls in this analysis (Table 8), likely because selection criteria for the broader research program required that the girl be unmarried, and because this paper further restricts the sample to girls who were also enrolled in school at baseline.[10]

Table 8: Age distribution for baseline schoolgirls

Age	N	Percent	Cumulative percent
13	341	15.04	15.04
14	397	17.50	32.54
15	525	23.15	55.69
16	389	17.15	72.84
17	250	11.02	83.86
18	179	7.89	91.75
19	99	4.37	96.12
20	58	2.56	98.68
21	21	0.93	99.60
22	9	0.40	100.00

Carrying out the same regressions only with girls aged 13–18 leads to results on self-reported school enrollment, marriage and pregnancy that are qualitatively similar as those presented in the original paper.[11] Girls receiving the UCT are more likely to report being enrolled in school, relative to girls in the control or CCT groups, and less likely to be married or pregnant by round 3 (Tables 9 and 10). In most cases, the coefficients in the school enrollment regressions from this younger sample are smaller than those using all girls aged 13–22, but this difference is not statistically significant.

[10] When considering the full sample, which includes girls who had already dropped out of school at baseline, the share of girls 19 and older is closer to 15 percent.

[11] All regressions in the measurement and estimation analysis also include controls for household mobile phone ownership and whether the household is female-headed, in addition to the controls used in the original paper.

Table 9: Age restricted to 13–18 – self-reported enrollment

	(1) 2008 Term 1	(2) 2008 Term 2	(3) 2008 Term 3	(4) 2009 Term 1	(5) 2009 Term 2	(6) 2009 Term 3	(7) TOTAL number of terms	(8) 2010 Term 1
= 1 if conditional schoolgirl	0.006	0.014	0.034*	0.039**	0.049***	0.055***	0.200***	–0.010
	(0.011)	(0.011)	(0.018)	(0.018)	(0.018)	(0.019)	(0.072)	(0.026)
= 1 if unconditional schoolgirl	0.029***	0.038***	0.037**	0.064***	0.088***	0.091***	0.348***	0.064***
	(0.010)	(0.010)	(0.017)	(0.022)	(0.022)	(0.021)	(0.078)	(0.022)
Number of observations	1,913	1,913	1,912	1,913	1,913	1,913	1,913	1,913

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

The statistically significant coefficients for the regressions on marriage and pregnancy are very similar when using the younger sample, as compared to those from the full sample.[12]

Table 10: Age restricted to 13–18 – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.010	–0.003	0.009	0.045*
	(0.013)	(0.023)	(0.014)	(0.026)
= 1 if unconditional schoolgirl	–0.022*	–0.072***	–0.008	–0.069***
	(0.012)	(0.022)	(0.016)	(0.022)
Number of observations	1,913	1,913	1,912	1,913

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

---

[12] In response to a reviewer comment, we also calculate (Westfall–Young) adjusted p-values to account for the family-wise error rate across these four outcomes in this particular analysis. Treating the outcomes in Table 10 as a family and focusing exclusively on the UCT treatment, for example, renders the following adjusted p-values for the UCT row: 0.1952, 0.0146, 0.5943 and 0.0194. This shifts the coefficient on the UCT in the first regression from p < 0.10 to larger and bumps the other two significant ones from p < 0.01 to p < 0.05. Nevertheless, the interpretation of the results is largely unchanged with this adjustment, as they still indicate that the UCT has a statistically significant negative effect, at conventional levels, on the likelihood of marriage and pregnancy by the end of the two-year study period. The adjusted p-values for these respective outcomes, using the full sample of baseline schoolgirls (Table 4, a replication of Table VII in the paper) for the UCT treatment only, are 0.0812, 0.0086, 0.678 and 0.026. Again, the UCT has a negative and significant impact, at conventional levels, on the likelihood of pregnancy and marriage by the end of the study.

Separately, we also consider heterogeneous effects by age as in the original paper but use a cut-off age of 18 rather than 15 given the reasoning above (Table 11). The effect of the UCT on marriage and pregnancy does not appear to be differentiated by older and younger girls, though the magnitude of these coefficients are somewhat larger when controlling for this higher age cut-off (7.1 and 6.8 percentage points, respectively, and in contrast to 5.1 and 5.9 percentage points).

Table 11: Heterogeneous effects by age with cut-off at 18

	(1) Ever married	**(2)** Ever pregnant
= 1 if conditional schoolgirl	–0.006 (0.025)	0.039 (0.027)
= 1 if unconditional schoolgirl	–0.071*** (0.021)	–0.068*** (0.022)
Conditional treatment * above 18 years old	–0.050 (0.093)	–0.094 (0.104)
Unconditional treatment * above 18 years old	–0.076 (0.079)	–0.032 (0.150)
= 1 if above 18 years old	–0.009 (0.042)	0.083* (0.044)
Number of observations	2,076	2,076

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Also, when restricting the age, the estimated effects of the cash transfer amounts versus the individual transfer do change, rendering the individual transfer significant for decreasing marriages and the household transfer significant for decreasing the likelihood of pregnancy (Table 12). In particular, the household transfer amount has a significant effect in decreasing the likelihood of marriage and pregnancy among UCT girls, with each dollar transferred decreasing the likelihood of marriage by 1.8 percentage points and the likelihood of pregnancy by 1.4 percentage points. The individual amount of the UCT transfer further decreases the likelihood of marriage by 2.5 percentage points per dollar.

Table 12: Age restricted to 13–18 – transfer amounts

	(1) Ever married	**(2)** Ever pregnant
Conditional treatment, individual amount	–0.005 (0.008)	0.000 (0.015)
Unconditional treatment, individual amount	–0.025** (0.011)	0.001 (0.013)
Conditional treatment, household amount	–0.001 (0.007)	0.004 (0.010)
Unconditional treatment, household amount	–0.018*** (0.007)	–0.014** (0.007)
Conditional treatment, minimum transfer	0.130* (0.073)	0.020 (0.063)
Unconditional treatment, minimum transfer	0.015 (0.070)	0.018 (0.093)
Number of observations	1,913	1,913

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Based on this analysis, it appears that a UCT targeting only school-aged girls would largely result in similar effects, as with the broader age range used in the original study, decreasing both marriage and pregnancy rates. Transfers made to parents decrease both marriage and pregnancy rates, whereas transfers made directly to the girls only decrease the likelihood of early marriage.

The next issue we consider is that older girls may respond differently to treatment not only if they are no longer of school age but also if they are not of the appropriate age for the grade they are due to attend. That is, in the same way that school-aged girls are less likely to respond to the treatment because they are more likely to attend school regardless, girls who have a history of attending school and thus making normal progress and enrolling in the right grade for their age might also be less likely to respond to the treatment.

To account for this, we consider the expected age of girls enrolled in each grade according to the Malawian school system, as detailed in the first two columns of Table 13. Allowing for up to two years of delays or repetition, we assign a maximum age for each grade that might still be considered normal school progress. Girls who report a given school grade but whose age is above the assigned cutoff are considered to be outside the appropriate age range for their grade. They are likely to be girls with weaker links to the school system, with more frequent dropouts, absences and/or repetitions. The two-year cutoff above the expected age is chosen arbitrarily but with the rationale that it allows for late starts in education, which is common in Malawi (World Bank 2010), and perhaps exceptional circumstances. Using a cutoff of one year yields qualitatively similar results.[13]

Table 13: Grade for age

Grade	Expected age	Maximum age
Primary 6	11/12	14
Primary 7	12/13	15
Primary 8	13/14	16
Secondary 1	14/15	17
Secondary 2	15/16	18
Secondary 3	16/17	19
Secondary 4	17/18	20

Note: For more information on the school system in Malawi and expected age for grade, see World Bank (2010).

It appears that this weaker attachment is very common in the study sample, as 50 percent of girls are older than the maximum age we assign for their grade (two years

[13] Setting the maximum age to just one year above the expected age for the grade lowers the share of girls in the “right” age for grade to 32 percent. Despite the smaller sample size, we find that the UCT has a statistically significant effect in decreasing marriage and pregnancy for this group. Similarly, for girls below the right grade for age, the UCT has an effect on decreasing marriage rates (11 percentage points by endline), but its coefficient is only significant at the 10 percent level in the regression using the endline pregnancy outcome as the dependent variable. Only 13 percent of girls in the study sample are in the expected age for their grade, so we do not conduct a separate analysis for this subsample.

above the expected age).[14] To account for this, we conduct the analysis separately for girls who are in the “right age for grade” (up to two years within range) and those older, although we note that this substantially lowers the power of our analysis for detecting true effects. Thus, coefficients on the UCT or CCT treatment variable that are found to not be statistically significant when using these split samples (as well as sub-samples based on baseline primary or secondary school enrollment, as done in the following sub-section) should be interpreted with caution.

The results for girls who are in the right age for grade are generally similar to those presented in the original study. The UCT has a greater impact on (self-reported) school enrollment, but the CCT has a positive impact as well (Table 14). The magnitude of the effect of the UCT in decreasing marriage by the final round is also similar, although its effect on pregnancy is greater than in the full sample (Table 15).

Table 14: Girls close to right grade for age – self-reported enrollment

	(1) 2008 Term 1	**(2)** 2008 Term 2	**(3)** 2008 Term 3	**(4)** 2009 Term 1	**(5)** 2009 Term 2	**(6)** 2009 Term 3	**(7)** TOTAL number of terms	**(8)** 2010 Term 1
= 1 if conditional schoolgirl	0.015	0.023**	0.036**	0.056***	0.051**	0.065***	0.248***	0.006
	(0.009)	(0.009)	(0.016)	(0.018)	(0.021)	(0.021)	(0.075)	(0.029)
= 1 if unconditional schoolgirl	0.024***	0.041***	0.041**	0.081***	0.086***	0.101***	0.377***	0.077***
	(0.009)	(0.008)	(0.017)	(0.021)	(0.021)	(0.023)	(0.079)	(0.027)
Number of observations	1,297	1,297	1,296	1,297	1,297	1,297	1,297	1,297

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 15: Girls close to right grade for age – marriage and pregnancy

	(1) Ever married R2	**(2)** Ever married R3	**(3)** Ever pregnant R2	**(4)** Ever pregnant R3
= 1 if conditional schoolgirl	0.007	0.004	–0.009	0.012
	(0.012)	(0.026)	(0.015)	(0.025)
= 1 if unconditional schoolgirl	–0.014	–0.060***	–0.017	–0.088***
	(0.009)	(0.022)	(0.018)	(0.025)
Number of observations	1,297	1,297	1,297	1,297

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

[14] This is consistent with high estimates of repetition from other sources. World Bank (2010) estimates that 16–38 percent of students in the eighth grade of primary school are repeaters; the higher end of the range accounts for those who may not have been held back by the school but who dropped out and returned to the same grade the next year or several years later.

An analysis using only girls who are too old for their reported grade suffers from a smaller sample but yields slightly different results. First, the effect of the CCT on (self-reported) school enrollment is very weak for this group, while the magnitude of the UCT effect is even larger than for the full sample (Table 16).

Table 16: Girls below right grade for age – self-reported enrollment

	(1) 2008 Term 1	**(2)** 2008 Term 2	**(3)** 2008 Term 3	**(4)** 2009 Term 1	**(5)** 2009 Term 2	**(6)** 2009 Term 3	**(7)** TOTAL number of terms	**(8)** 2010 Term 1
= 1 if conditional schoolgirl	–0.003	0.012	0.050*	0.039	0.063*	0.049	0.211*	0.007
	(0.023)	(0.023)	(0.029)	(0.031)	(0.033)	(0.034)	(0.126)	(0.050)
= 1 if unconditional schoolgirl	0.057***	0.073***	0.076**	0.063	0.128**	0.124**	0.520***	0.075
	(0.019)	(0.027)	(0.037)	(0.049)	(0.051)	(0.048)	(0.184)	(0.052)
Number of observations	779	779	779	779	779	779	779	779

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

At odds with this result, it seems that while the UCT decreases the likelihood of marriage among this group of girls who have weaker ties to school, the CCT may increase their likelihood of being pregnant by as much as 9.8 percentage points by the end of the program (Table 17). There is no statistically significant effect of the UCT on the likelihood of a girl ever being pregnant in this restricted sample. This may certainly be driven by the low statistical power of this analysis using this sub-sample – well below 80 percent – although it could also indicate that a UCT aiming to decrease pregnancies may be more effective if it focuses on girls with a strong attachment to school, who are close to the right age for their grade and do not have a history of repetitions or dropouts. If the program's goal is to prevent marriages among young girls, however, including girls who are below the right age for grade may actually lead to a higher average impact.

Table 17: Girls below right grade for age – marriage and pregnancy

	(1) Ever married R2	**(2)** Ever married R3	**(3)** Ever pregnant R2	**(4)** Ever pregnant R3
= 1 if conditional schoolgirl	0.003	–0.018	0.052*	0.098**
	(0.024)	(0.036)	(0.030)	(0.044)
= 1 if unconditional schoolgirl	–0.060**	–0.121***	0.011	–0.051
	(0.024)	(0.040)	(0.054)	(0.058)
Number of observations	779	779	778	779

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.1.2 Additional transfer for secondary schoolgirls in UCT arm

Next, we explore the potential role of a feature of the program design in driving the obtained results. The program treatment had a UCT and a CCT arm, both of which included transfers to the girl and to her parents at randomly assigned amounts. The transfer to the girls ranged from US$1–5, with a mean of US$3, while the randomly assigned amount to the parents ranged from US$4–10, with a mean of US$7.

However, costs imposed by the school system in Malawi created an additional complexity to the transfer, as primary school is generally free but there are non-negligible costs associated with secondary school attendance (Baird et al. 2011, footnote 18). To address this issue, the research team covered the school fees for secondary schoolgirls in the CCT arm and made an additional transfer of 1,000 kwacha (close to US$7) per month – the average school fee for public secondary schools in the sample – to all parents of girls in the UCT arm who were secondary schoolgirls at baseline.

Given that 47 percent of parents of girls in the UCT arm received this additional transfer, on average doubling the monthly amount received, we conduct additional analyses to investigate whether it may play a role in the observed effects.

The original paper notes this additional transfer but does not account for it in the main analysis, grouping all girls in each arm together, regardless of schooling level. It does, however, address this feature in Appendix E through two robustness checks. In the first, the authors run a set of regressions focusing only on girls enrolled in grades 7 or below at baseline and find results in the same direction as in the main paper. The magnitude of the effects on school attendance and marriage, though, are substantially smaller for these primary schoolgirls, compared to the full sample. For example, for primary schoolgirls, the CCT increases school attendance by 0.346 terms, whereas the average effect for the full sample is 0.535. We explore heterogeneous effects by baseline primary or secondary school status in more detail in this sub-section.

In Appendix E, the authors also explain that girls in the CCT arm who moved from primary to secondary school between the two treatment years received the same tuition coverage in the second year as girls who started off signed up for secondary school at the beginning of treatment. In contrast, girls in the UCT arm who completed primary school after the first year received only the base transfer the second year. They did not receive the additional 1,000 kwacha transfer, as UCT girls who started off as baseline secondary schoolgirls did. To overcome this design issue, the other robustness check in Appendix E excludes grade 8 girls from the analysis, while still keeping primary and secondary schoolgirls together in the analysis.

In this section, we discuss potential heterogeneous effects of this additional transfer, accounting for the unique experience with the program for baseline grade 8 girls in the UCT arm.[15]

---

[15] Baseline grade 8 girls in the UCT arm who did not complete grade 8 at the end of the first year continued to receive only the base transfer, as designed. The unique experience rests on baseline grade 8 girls who completed grade 8 but did not receive an additional transfer in their second year. Since we only have self-reported data, we cannot reliably distinguish between these

First, we note that there appears to be a small mismatch between highest grade at baseline and the households that are recorded as receiving the additional transfer (Table 18). We assume the record on transfers is accurate, as this was logged by the research team to ensure proper transfer amounts, whereas there may be some inaccuracies in the grade information from the baseline survey. As there is no reason to believe there would be a bias in the reported grade at baseline (since respondents did not know ex ante what the treatment would be, much less the differentiated transfer amounts for primary and secondary schoolgirls), we do not expect this discrepancy to have a meaningful effect on the overall results.

Table 18: UCT arm – number of girls by grade completed (reported at baseline) and additional transfers made

Baseline grade completed	Household did not receive additional transfer	Household received additional transfer
5	10	0
6	60	0
7	68	7
8	5	33
9	3	39
10	4	34
11	1	17
12	0	2
Total	151	132

In Tables 19 and 20, we present results for regressions that simply include a dummy variable for households that received the additional cash transfer (UCT girls who had completed grade 8 or higher at baseline) and do not include baseline grade 8 girls when considering 2009–2010 outcomes (columns 4–8). Although the effect of the UCT remains positive and statistically significant, the coefficients generally decrease when this variable is included. The UCT appears to increase the number of terms completed during the two-year period by 0.219, but girls in households receiving the additional transfer completed, on average, an extra 0.444 terms more than the control group (Table 19).

two groups of girls, and therefore argue for excluding all baseline grade 8 girls for round 2 outcomes in our analysis in this section. Although this experience only applies to UCT girls, we drop grade 8 girls across all arms, as we do not know about interactions between girls at school and are concerned about UCT girls in grade 8 knowing that their classmates who completed primary school with them but are in the CCT arm are receiving tuition coverage, or, alternatively, girls in CCT arm knowing that some girls receive the additional transfer and some do not.

Table 19: Additional transfer dummy – self-reported enrollment

	(1) 2008 term 1	(2) 2008 term 2	(3) 2008 term 3	(4) 2009 term 1	(5) 2009 term 2	(6) 2009 term 3	(7) TOTAL number of terms	(8) 2010 term 1
= 1 if conditional schoolgirl	0.006	0.018*	0.039**	0.048***	0.053***	0.060***	0.226***	0.006
	(0.011)	(0.011)	(0.017)	(0.017)	(0.019)	(0.019)	(0.071)	(0.026)
= 1 if unconditional schoolgirl	0.015*	0.021*	0.010	0.038	0.065**	0.078***	0.199**	0.086**
	(0.009)	(0.011)	(0.019)	(0.024)	(0.027)	(0.028)	(0.094)	(0.043)
Household received additional transfer	0.041***	0.061***	0.095***	0.079**	0.070*	0.085**	0.464***	–0.044
	(0.015)	(0.017)	(0.026)	(0.036)	(0.036)	(0.037)	(0.139)	(0.081)
Number of observations	2,076	2,076	2,075	2,012	2,012	2,012	2,012	2,012

Note: Columns 4–7 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

More markedly, the effect of the UCT alone on marriage disappears when accounting for the additional transfer in this manner. That is, girls in households receiving the additional transfer are 9 percentage points less likely to be married by round 3, but the coefficient is not significant for girls not receiving that extra transfer (Table 20). In contrast, the effect of the UCT on decreasing pregnancy becomes insignificant by the end of two years.

Table 20: Additional transfer dummy – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.007	–0.004	0.014	0.040
	(0.012)	(0.024)	(0.014)	(0.027)
= 1 if unconditional schoolgirl	–0.010	–0.043	0.009	–0.051
	(0.012)	(0.027)	(0.016)	(0.034)
Household received additional transfer	–0.038***	–0.092***	–0.036	–0.042
	(0.014)	(0.041)	(0.031)	(0.045)
Number of observations	2,076	2,012	2,075	2,012

Note: Columns 2 and 4 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Another approach for taking this additional transfer into account is to include it in the household amount when measuring the effects of the amounts and minimum transfer. Doing so changes the range of the unconditional treatment household amount variable to US$4–17.14 and the mean to US$10.1 (as households whose daughters are in primary school are included in this computation but did not receive the additional transfer). The effect of each dollar transferred to the household on marriage does not meaningfully change with this modification, however, as each dollar transferred to the household decreases the likelihood of marriage by round 3 by 1.5 percentage points (Table 21).

Table 21: Accounting for additional transfer – transfer amounts

	(1) Ever married	(2) Ever pregnant
Conditional treatment, individual amount	–0.002	0.001
	(0.008)	(0.014)
Unconditional treatment, individual amount	–0.018	0.014
	(0.014)	(0.015)
Conditional treatment, household amount	0.000	0.006
	(0.007)	(0.011)
Unconditional treatment, household amount	–0.015*	–0.003
	(0.008)	(0.013)
Conditional treatment, minimum transfer	0.061	–0.100
	(0.081)	(0.107)
Unconditional treatment, minimum transfer	–0.001	–0.006
	(0.072)	(0.097)
Number of observations	2,012	2,012

Note: Excludes girls whose highest grade completed at baseline was grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A complication with the additional transfer is that it is not randomly assigned. Rather, it is given to all girls whose current grade at baseline would be in secondary school. As such, a dummy for the additional transfer is also an indicator that the girl is in secondary school and in the UCT arm. To separate these two factors, we run additional regressions separately for primary and secondary schoolgirls. In this sense, the regressions for primary schoolgirls compare the UCTs and CCTs of the same amount to the control group, while the regressions for secondary schoolgirls are actually comparing a CCT of mean US$7 plus school fee coverage and a UCT of mean US$14 to the control group.[16]

Again, given the decrease in power when using just the primary or secondary schoolgirls' sub-samples, we caution against giving much importance to coefficients that are not statistically significant here.

[16] We exclude baseline grade 8 girls for the analysis on secondary schoolgirls for 2009 and round 3 outcomes, as they are not comparable to other primary schoolgirls in the UCT arm, who did not have to pay for school fees throughout, or to baseline secondary schoolgirls in the UCT arm, whose households received the additional US$7 each month. If they made normal progress through school, they would be eligible for secondary school in the middle of the study, after the round 2 surveys.

Nevertheless, for girls in primary school, the effect of the UCT on (self-reported) school enrollment is slightly lower than with the full sample, although the effect of the CCT remains similar, raising attendance by 0.21 terms over the two years (Table 22). The effect of these transfers on marriage are also qualitatively similar to the full sample (Table 24). However, the UCT decreases the chance of pregnancy for primary schoolgirls by 9.2 percentage points, in contrast to 6.7 percentage points found in the original study for the full sample. That is, the estimated effect of the UCT in deterring pregnancy is even larger among girls in primary school, compared to the control group and the CCT group, than when using the full sample (Table 24).

Table 22: Baseline primary schoolgirls – self-reported enrollment

	(1) 2008 term 1	**(2)** 2008 term 2	**(3)** 2008 term 3	**(4)** 2009 term 1	**(5)** 2009 term 2	**(6)** 2009 term 3	**(7)** TOTAL number of terms	**(8)** 2010 term 1
= 1 if conditional schoolgirl	0.006	0.016	0.046**	0.037**	0.050**	0.054**	0.210***	0.028
	(0.012)	(0.011)	(0.018)	(0.017)	(0.020)	(0.022)	(0.075)	(0.031)
= 1 if unconditional schoolgirl	0.021**	0.028**	0.019	0.057**	0.094***	0.092***	0.311***	0.105***
	(0.009)	(0.011)	(0.020)	(0.023)	(0.024)	(0.022)	(0.078)	(0.024)
Number of observations	1,490	1,490	1,490	1,490	1,490	1,490	1,490	1,490

Note: Columns 4–7 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 23: Baseline secondary schoolgirls – self-reported enrollment

	(1) 2008 term 1	**(2)** 2008 term 2	**(3)** 2008 term 3	**(4)** 2009 term 1	**(5)** 2009 term 2	**(6)** 2009 term 3	**(7)** TOTAL number of terms	**(8)** 2010 term 1
= 1 if conditional schoolgirl	0.011	0.029	0.027	0.089*	0.059	0.074*	0.294*	–0.051
	(0.022)	(0.029)	(0.038)	(0.047)	(0.038)	(0.042)	(0.160)	(0.045)
= 1 if unconditional schoolgirl	0.051***	0.090***	0.117***	0.104**	0.105**	0.133**	0.608***	0.006
	(0.019)	(0.023)	(0.028)	(0.051)	(0.052)	(0.054)	(0.189)	(0.070)
Number of observations	586	586	585	586	586	586	586	586

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Also, the estimated effect of the UCT on (self-reported) school attendance is larger when analysis is restricted to secondary schoolgirls who are receiving a larger direct transfer amount than those complying with the CCT treatment (Table 23). This larger UCT appears to deter marriage by 9.5 percentage points among secondary schoolgirls (Table 25), and each dollar transferred decreases the likelihood of marriage by 1.9 percentage points. In contrast, we do not observe any significant effect on pregnancy, although the lack of power for this test (well below 80 percent) prevents us from drawing any conclusions in this respect (Table 25).

Table 24: Baseline primary schoolgirls – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever married R3	(4) Ever pregnant R2	(5) Ever pregnant R3	(6) Ever pregnant R3
= 1 if conditional schoolgirl	–0.004	–0.031	–0.030	–0.000	0.020	0.037
	(0.012)	(0.020)	(0.021)	(0.016)	(0.032)	(0.031)
= 1 if unconditional schoolgirl	–0.024**	–0.070***	–0.066**	–0.011	–0.092***	–0.085***
	(0.011)	(0.024)	(0.028)	(0.017)	(0.024)	(0.029)
Number of observations	1,490	1,490	1,206	1,489	1,490	1,206

Note: Columns 3 and 6 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 25: Baseline secondary schoolgirls – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.035	0.077	0.044	0.096*
	(0.031)	(0.054)	(0.047)	(0.050)
= 1 if unconditional schoolgirl	–0.029	–0.095**	0.011	–0.016
	(0.020)	(0.045)	(0.038)	(0.043)
Number of observations	586	586	586	586

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Again, we observe a possible impact of the CCT in increasing the likelihood of pregnancy among baseline secondary schoolgirls by the end of the study, significant at the 10 percent level.

These findings by baseline education level are similar to those on school attachment: a policy aiming to discourage pregnancies will be more effective if it focuses on girls in primary school, but including secondary schoolgirls will have an effect in decreasing marriage rates. Also, there is some evidence that CCTs made to secondary schoolgirls and to girls with a weak attachment to school might actually increase their likelihood of pregnancy (Table 25).

Tables 26 and 27 present results for regressions, distinguishing between parent and individual amounts. We do not observe any consistent pattern for baseline primary schoolgirls; for secondary schoolgirls, we find a negative and statistically significant coefficient for the UCT household amount when considering the likelihood of marriage, but positive coefficient on the individual amount of the UCT when considering pregnancy.

Table 26: Baseline primary schoolgirls – transfer amounts

	(1) Ever married	(2) Ever married	(3) Ever pregnant	(4) Ever pregnant
Conditional treatment, individual amount	0.004	–0.003	0.003	–0.006
	(0.010)	(0.012)	(0.020)	(0.023)
Unconditional treatment, individual amount	–0.027*	–0.022	–0.011	–0.016
	(0.015)	(0.016)	(0.014)	(0.021)
Conditional treatment, household amount	0.001	0.002	0.004	0.003
	(0.007)	(0.007)	(0.012)	(0.012)
Unconditional treatment, household amount	–0.007	–0.018	–0.005	–0.018*
	(0.009)	(0.011)	(0.009)	(0.011)
Conditional treatment, minimum transfer	–0.049	–0.035	–0.019	0.034
	(0.058)	(0.066)	(0.114)	(0.120)
Unconditional treatment, minimum transfer	0.057	0.124	–0.028	0.083
	(0.092)	(0.105)	(0.084)	(0.108)
Number of observations	1,490	1,206	1,490	1,206

Note: Columns 2 and 4 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 27: Baseline secondary schoolgirls – transfer amounts

	(1) Ever married	(2) Ever pregnant
Conditional treatment, individual amount	–0.006	–0.000
	(0.020)	(0.025)
Unconditional treatment, individual amount	0.007	0.056**
	(0.022)	(0.028)
Conditional treatment, household amount	0.005	0.026
	(0.023)	(0.022)
Unconditional treatment, household amount	–0.019***	–0.006
	(0.007)	(0.007)
Conditional treatment, minimum transfer	0.149	–0.106
	(0.117)	(0.119)
Unconditional treatment, minimum transfer	0.060	–0.084
	(0.180)	(0.185)
Number of observations	586	586

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.1.3 Additional effects on treated girls, by treatment saturation

A final consideration in this section pertains to another detail in the design of the program. In the UCT and CCT EAs, not all girls were selected for the transfer. Rather, just as each area was randomly assigned into one of the treatments or the control group, it was also assigned for one-third, two-thirds or all of the schoolgirls in the area to receive the treatment offer (Baird et al. 2011, footnote 15). We test here whether these shares had an effect on the outcomes of interest for the treated girls. Specifically, we check whether, aside from the direct effect from the transfer offer, there was also an indirect effect, as girls were more (or less) likely to be enrolled in school (self-reported), or get married or become pregnant, depending on whether a larger share of girls in their area were receiving the same offer.

Table 28 presents our results for self-reported school enrollment, where the first two coefficients in each column represent those for the CCT and UCT dummies, respectively. The subsequent coefficients represent the additional effect that may come from being in an EA where 66 percent or 100 percent of eligible girls were assigned the same treatment.[17] We find only weak evidence of additional effect on self-reported school enrollment for treated girls, as girls in UCT areas with a high concentration of other recipients are less likely to report attending school during some terms. Given the argument in the original study that self-reported enrollment is unreliable, it is not clear whether this should be interpreted as the higher concentration attenuating the direct positive effect of the UCT on attendance rates or a higher concentration decreasing the pressure to over-report school attendance.

There does not appear to be any additional effect from the saturation of transfers on the likelihood of marriage or pregnancy (Table 29) for treated girls, as the effect is concentrated on individual assignment into the UCT arm.[18]

[17] The variation in the saturation was assigned at the EA level, not the school level. We do not know whether there are several schools per EA or, on the other extreme, one school for many EAs. Understandably, school information has been removed from the World Bank data set (it is personally identifiable information), and we have not received a response from the original authors regarding the catchment areas of schools in this setting.

[18] The analyses for additional effects from EA treatment saturation on pregnancy and marriage outcomes by the endline for treated girls are powered at more than 75 percent.

Table 28: Additional effects on treated girls – self-reported enrollment

	(1) 2008 term 1	(2) 2008 term 2	(3) 2008 term 3	(4) 2009 term 1	(5) 2009 term 2	(6) 2009 term 3	(7) TOTAL number of terms	(8) 2010 term 1
= 1 if conditional schoolgirl	0.024*	0.042**	0.068**	0.068**	0.053*	0.037	0.294**	–0.027
	(0.014)	(0.019)	(0.030)	(0.034)	(0.032)	(0.037)	(0.126)	(0.044)
= 1 if unconditional schoolgirl	0.044***	0.066***	0.063***	0.108***	0.126***	0.141***	0.552***	0.093**
	(0.010)	(0.009)	(0.023)	(0.035)	(0.041)	(0.036)	(0.130)	(0.044)
Conditional schoolgirl; 66% offers	–0.040**	–0.040*	–0.048	–0.033	0.012	0.046	–0.104	0.069
	(0.018)	(0.022)	(0.029)	(0.038)	(0.038)	(0.041)	(0.148)	(0.058)
Conditional schoolgirl; 100% offers	–0.011	–0.028	–0.032	–0.021	–0.010	0.023	–0.079	0.026
	(0.015)	(0.021)	(0.030)	(0.035)	(0.038)	(0.043)	(0.132)	(0.055)
Unconditional schoolgirl; 66% offers	–0.012	–0.025	–0.027	–0.083**	–0.051	–0.058	–0.257**	–0.032
	(0.012)	(0.016)	(0.033)	(0.038)	(0.043)	(0.040)	(0.130)	(0.049)
Unconditional schoolgirl; 100% offers	–0.027**	–0.039**	–0.016	–0.044	–0.046	–0.068*	–0.239*	–0.022
	(0.013)	(0.018)	(0.024)	(0.042)	(0.046)	(0.040)	(0.142)	(0.056)
Number of observations	2,076	2,076	2,075	2,076	2,076	2,076	2,076	2,076

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 29: Additional effects on treated girls – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.010	0.021	–0.001	0.087*
	(0.029)	(0.055)	(0.025)	(0.051)
= 1 if unconditional schoolgirl	–0.020	–0.106***	–0.017	–0.092**
	(0.021)	(0.030)	(0.029)	(0.039)
Conditional schoolgirl; 66% offers	–0.004	–0.060	0.048	–0.077
	(0.031)	(0.058)	(0.030)	(0.063)
Conditional schoolgirl; 100% offers	–0.006	–0.017	–0.005	–0.059
	(0.029)	(0.056)	(0.027)	(0.061)
Unconditional schoolgirl; 66% offers	–0.015	0.035	0.002	0.014
	(0.023)	(0.039)	(0.035)	(0.049)
Unconditional schoolgirl; 100% offers	–0.012	0.064	0.033	0.061
	(0.021)	(0.042)	(0.036)	(0.043)
Number of observations	2,076	2,076	2,075	2,076

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.2 Additional analysis

The survey instruments used in this study are very rich, and while working on the planned replication with this data, we found additional variables we believed were worth exploring. In this section, we discuss additional exploratory analysis we conducted after becoming familiar with the data during the replication process. First, we cover what is more commonly understood as a spillover effect – the impact on girls not given the transfer directly but who live in areas where transfers to schoolgirls were made. Next, we conduct an analysis with an alternative outcome of interest with respect to (self-reported) school attendance, given the condition of the CCT. In the final two sub-sections here, we look at baseline variables relevant to the main outcomes studied in this replication exercise and the effect of including them in our analysis.

4.2.1 Spillover effects on non-treated girls

In the previous section, we showed that, for treated girls, no additional effect from living in a village where more (or fewer) girls also received the CCT or UCT was identified. Here, we cover another kind of spillover effect – that on schoolgirls who were eligible for the transfer and lived in treated EAs but were nonetheless not selected to receive the transfers during the two years. These girls are not included in the rest of the analysis in this paper, but, to explore the potential impact on this group of girls, we compare baseline schoolgirls who did not receive the transfers but lived in EAs with 33 percent or 66 percent saturation of schoolgirls receiving either the UCT or CCT to schoolgirls in control EAs.[19] None of the girls in this investigation received a transfer; they only differ by the extent to which other schoolgirls in the same EA received the transfer, from 0 percent (control) to 33 percent and 66 percent.

We generally do not observe any significant spillover effect on the (self-reported) school enrollment of non-treated girls (Table 30).

[19] This sub-section was added in response to reviewer comments, and the analysis included herein is exploratory. We do not study non-treated girls living treated EAs ("spillover girls") in more detail in our analysis.

Table 30: Spillover effects on non-treated girls – self-reported enrollment

	(1) 2008 term 1	(2) 2008 term 2	(3) 2008 term 3	(4) 2009 term 1	(5) 2009 term 2	(6) 2009 term 3	(7) TOTAL number of terms	(8) 2010 term 1
33% CCT offers	0.015	–0.012	–0.005	–0.006	0.022	0.011	0.026	0.010
	(0.015)	(0.019)	(0.028)	(0.027)	(0.025)	(0.028)	(0.108)	(0.031)
66% CCT offers	–0.027	–0.039	–0.063	0.010	0.034	0.059*	–0.025	0.020
	(0.027)	(0.027)	(0.060)	(0.039)	(0.035)	(0.031)	(0.187)	(0.055)
33% UCT offers	0.003	–0.006	–0.044	0.011	–0.013	–0.013	–0.060	0.036
	(0.014)	(0.020)	(0.053)	(0.048)	(0.048)	(0.043)	(0.199)	(0.038)
66% UCT offers	–0.014	–0.031	0.010	0.038	0.090**	0.076*	0.171	–0.046
	(0.044)	(0.036)	(0.042)	(0.062)	(0.044)	(0.045)	(0.259)	(0.050)
Number of observations	1,724	1,724	1,723	1,724	1,724	1,724	1,724	1,724

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In contrast, we observe spillover effects, in both their marriage and pregnancy outcomes, of non-treated girls living in UCT EAs. By the end of the study, girls living in EAs where other girls received unconditional transfers were 7–10 percentage points less likely to be married and 8–9 percentage points less likely to have ever been pregnant, with no statistically significant difference in these spillover effects between those in EAs with 33 percent saturation and with 66 percent saturation (Table 31). These effects are larger in magnitude than the effects on the UCT-recipient girls themselves (7.9 for marriage and 6.7 for pregnancy), although not significantly so. There is no such effect on girls in CCT villages by the end of the study.

Table 31: Spillover effects on non-treated girls – marriage and pregnancy

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
33% CCT offers	0.012	0.024	–0.001	0.085*
	(0.028)	(0.055)	(0.025)	(0.051)
66% CCT offers	0.007	–0.036	0.046**	0.008
	(0.016)	(0.025)	(0.020)	(0.041)
33% UCT offers	–0.018	–0.104***	–0.018	–0.094**
	(0.021)	(0.029)	(0.028)	(0.038)
66% UCT offers	–0.034***	–0.070**	–0.015	–0.081**
	(0.012)	(0.029)	(0.022)	(0.033)
Number of observations	2,076	2,076	2,075	2,076

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Although these results are exploratory, they may represent a true spillover, whereby girls whose friends were less likely to be married or pregnant due to the treatment also became less likely to have either of these two outcomes, due to peer pressure or influences from friends or family. Alternatively, this could point to the salience of the program and its eligibility: as non-treated girls knew about the program and that other never married girls in their EA were receiving the UCT, they or their families might have avoided marriage to retain eligibility in case the program was expanded. This does not explain the substantially lower pregnancy rates, however, as never having been pregnant was not an eligibility criterion, and people seem to have understood that, as discussed above (sections 4.1.1 and 4.1.2) with respect to the impacts of the UCT on older girls and those with a weaker school attachment.

4.2.2 School attendance – at least 80 percent

We now return to the girls who actually received the UCT or CCT transfers – the group of interest in *Cash or condition*. Although we do not have teacher-reported information, to grapple a bit more with the measured effects on school attendance, we consider an alternative self-reported outcome for which the answer could be more salient for CCT girls. The measurement used for self-reported school enrollment in the original paper is based on a question in the survey instrument, asking whether the respondent "attended school any time during [term]." The CCT condition, however, was attendance of at least 80 percent of the days in a given month.

Here, we briefly explore the effect of the treatments on self-reported attendance of at least this level, using the following question from the third survey round: "During a typical week in [term], did you attend at least four out of every five days school was in session?"[20] CCT girls may have more accurate answers to this question than the other two arms, as they may recall whether they received payments during those terms. It is unclear, however, whether girls in the other two arms would potentially over-report or under-report this result, particularly in relation to the more general question used in the paper.

Nevertheless, the results are qualitatively similar to those using the first outcome: both the CCT and the UCT have a large positive impact in raising self-reported high attendance. The CCT increased attendance during the conditional period, which lasted until the final term of 2009, although this effect disappears once the program ends. In contrast, girls in the UCT arm report higher attendance rates even once the transfers have ended. This may offer some evidence of the power of the conditionality in encouraging attendance, as girls in the CCT arm attend at higher levels (self-reported) for only as long as they are receiving an incentive to do so.

The higher attendance rates among UCT girls in 2010 might be due to over-reporting, although it is unclear why they would do so after the program has closed. Rather, it could be that both treatments encourage school attendance (self-reported), but the CCT leads to a drop, once the conditionality disappears, that the UCT can avoid by not explicitly highlighting the association of the transfers to school attendance or fees. Of course, this interpretation assumes the accuracy of self-reported attendance information.

[20] This question is not included in the second round of the survey.

Table 32: Self-reported school attendance – at least 4 out of 5 days

	(1) 2009 term 1	(2) 2009 term 2	(3) 2009 term 3	(4) 2010 term 1
= 1 if conditional schoolgirl	0.058**	0.079***	0.089***	0.031
	(0.023)	(0.021)	(0.024)	(0.030)
= 1 if unconditional schoolgirl	0.090***	0.106***	0.116***	0.105***
	(0.028)	(0.028)	(0.030)	(0.031)
Number of observations	2,061	2,066	2,065	1,966
Prob > F (conditional = unconditional)	0.3126	0.3459	0.3910	0.0434

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2.3 Differences in other baseline variables

Next, we discuss other baseline variables we identified as potentially relevant for the main outcomes studied in this replication exercise, as they touched on the same topics as the outcome variables focused on here (e.g. aspirations for education and marriage). These were found through a thorough reading of the baseline instruments (as was the question above on school attendance of at least 4 out of 5 days) and analyzed first for their distribution at baseline. They generally relate to the girls' own earnings, their plans and aspirations for education and marriage, and exposure to risky sex, as presented in Table 33.

When asked about any income they earned over the last 12 months, for example, girls assigned to the UCT arm report on average lower individual earnings than girls in the CCT arm or control group. And although only about half of the girls have any positive income over the previous 12 months, this share is closer to one-third among girls in the UCT arm.

Girls in the UCT arm are also more likely to say they want to get married at some point and more likely to report that their closest (unmarried female) friend has had sex in the last 12 months.[21] The latter may be indicative of girls who are more exposed to sex, perceive sex outside of marriage as more common and/or are more likely to be close to an unmarried young woman who is pregnant or has children. Girls in the UCT arm also tend to report a higher likelihood that they will be in school next year than the control group, as well as a higher desired level of schooling before they get married. Given these differences across arms, we re-run the regressions with controls for these variables.

Table 33: Additional variables – baseline means and balance

	(1) Control group mean	(2) = 1 if conditional schoolgirl	(3) = 1 if unconditional schoolgirl	(4) p-value (conditional-unconditional)
Individual-level variables				

[21] For this latter measurement, we include girls who *suspect* that their closest (unmarried female) friend has had sex in the last 12 months.

	(1) Control group mean	(2) = 1 if conditional schoolgirl	(3) = 1 if unconditional schoolgirl	(4) p-value (conditional-unconditional)
Individual earnings	667.20	836.84	349.27***	0.001
	(1644.79)	(2071.39)	(781.64)	
Non-zero individual earnings	0.507	0.543	0.376*	0.051
	(0.500)	(0.499)	(0.485)	
Household school expenditures	3482.61	3202.53	3937.18	0.121
	(5289.36)	(4581.50)	(5557.46)	
Likelihood in school next year +	3.882	3.888	3.940*	0.153
	(0.410)	(0.472)	(0.319)	
Likelihood complete secondary +	3.778	3.742	3.844	0.056
	(0.510)	(0.533)	(0.448)	
Ever marriage promise	0.064	0.067	0.089	0.265
	(0.245)	(0.250)	(0.285)	
Want to get married	0.870	0.843	0.914	0.020
	(0.337)	(0.364)	(0.281)	
Plan on marriage in 3 years	0.006	0.017*	0.025	0.657
	(0.076)	(0.131)	(0.158)	
Desired schooling +	5.608	5.546	5.708**	0.056
	(0.539)	(0.672)	(0.456)	
Kids timing +	2.542	2.533	2.399	0.259
	(1.025)	(1.061)	(1.003)	
Best friend had sex	0.326	0.272*	0.376	0.044
	(0.469)	(0.445)	(0.485)	
Worry about AIDS +	1.525	1.468	1.600	0.166
	(0.691)	(0.714)	(0.696)	
Observations	1285–1490	424–502	247–281	

Note: + indicates an ordinal variable. Likelihood is assigned (1) no likelihood, (2) low, (3) medium and (4) high likelihood. The amount of schooling a girl would like to have before marriage (desired schooling) is assigned values 1–6 for responses ranging from no schooling to tertiary education. The desired interval between marriage and having kids ranges from (1) less than a year to (5) four years or more. Responses for whether a girl worries about catching HIV/AIDS are categorized as (1) not worried at all, (2) worries a little and (3) worries a lot. Only girls who reported wanting to get married at some point were asked about their desired schooling level before marriage and the desired timing of children after marriage.

Among these baseline variables, having some earnings of one's own, as well as having a best friend who had sex in the 12 months prior to baseline, appear to be related to the outcomes of interest. In particular, if a girl reports at baseline that she believes her best friend had sex, she is less likely to report being enrolled in school during the study, and reports attending 0.242 fewer terms over the study period (Table 34). She is also 6.4 percentage points more likely to have been pregnant by the end of the study, a coefficient of similar magnitude (but opposite direction) as that of the UCT transfer (Table 35). In contrast, girls who have some individual earnings over the 12 months prior to the survey are 4.2 percentage points more likely to be married by the end of two years but not more likely to have ever been pregnant. The coefficients on the treatment variables, however, are essentially the same as in the original paper when including these baseline variables in the regressions for self-reported enrollment, marriage and pregnancy.

Table 34: Additional variables – self-reported enrollment

	(1) 2008 term 1	**(2)** 2008 term 2	**(3)** 2008 term 3	**(4)** 2009 term 1	**(5)** 2009 term 2	**(6)** 2009 term 3	**(7)** TOTAL number of terms	**(8)** 2010 term 1
= 1 if conditional schoolgirl	0.003	0.020*	0.032**	0.047***	0.053***	0.062***	0.218***	0.011
	(0.010)	(0.010)	(0.015)	(0.017)	(0.019)	(0.019)	(0.068)	(0.028)
= 1 if unconditional schoolgirl	0.034***	0.045***	0.047***	0.076***	0.103***	0.108***	0.413***	0.070***
	(0.009)	(0.010)	(0.016)	(0.023)	(0.024)	(0.023)	(0.082)	(0.025)
Non-zero individual earnings	0.009	–0.012	0.002	–0.013	–0.001	–0.007	–0.022	–0.021
	(0.009)	(0.011)	(0.014)	(0.017)	(0.020)	(0.021)	(0.073)	(0.028)
Best friend had sex	–0.012	–0.006	–0.046***	–0.052**	–0.075***	–0.052**	–0.242***	–0.082**
	(0.011)	(0.011)	(0.014)	(0.022)	(0.026)	(0.023)	(0.078)	(0.034)
Number of observations	2,042	2,042	2,042	2,042	2,042	2,042	2,042	2,042

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 35: Additional variables – marriage and pregnancy

	(1) Ever married R2	**(2)** Ever married R3	**(3)** Ever pregnant R2	**(4)** Ever pregnant R3
= 1 if conditional schoolgirl	0.005	–0.011	0.018	0.042
	(0.013)	(0.023)	(0.015)	(0.028)
= 1 if unconditional schoolgirl	–0.027**	–0.074***	–0.008	–0.068***
	(0.011)	(0.022)	(0.018)	(0.025)
Non-zero individual earnings	0.030***	0.042**	–0.005	–0.017
	(0.009)	(0.018)	(0.014)	(0.025)
Best friend had sex	0.018	0.014	0.036**	0.064**
	(0.011)	(0.022)	(0.016)	(0.030)
Number of observations	2,042	2,042	2,041	2,042

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Next, we re-run the regression exploring heterogeneous effects by age using our preferred specification, which employs age 18 as the cutoff for the age dummy, and including the additional baseline variables (Table 36). Again, the coefficients on the treatment are slightly larger, but qualitatively similar to that presented in the paper, with UCTs decreasing the likelihood of marriage by 7 percentage points on average and pregnancy by 6.8 percentage points on average. There does not appear to be a differentiated effect by age. Girls who have some earnings at baseline are nonetheless more likely to be married by the end of the study period, whereas those who believe their close friend had sex within the year prior to baseline are 6.7 percentage points more likely to get pregnant by the third survey round.

Table 36: Additional variables – age heterogeneity

	(1) Ever married	**(2)** Ever pregnant
= 1 if conditional schoolgirl	–0.008	0.050*
	(0.023)	(0.027)
= 1 if unconditional schoolgirl	–0.070***	–0.068***
	(0.022)	(0.024)
Conditional treatment *above 18 years old	–0.050	–0.135
	(0.097)	(0.120)
Unconditional treatment * above 18 years old	–0.055	0.007
	(0.082)	(0.152)
= 1 if above 18 years old	–0.219	–0.399
	(0.654)	(0.794)
Non-zero individual earnings	0.041**	–0.018
	(0.019)	(0.025)
Best friend had sex	0.014	0.067**
	(0.022)	(0.030)
Number of observations	2,042	2,042

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Lastly, as in the previous sub-section, we run the regressions again separately for girls enrolled in primary and in secondary school, as the transfer amounts under the UCT, as well as the total value of the benefits for compliant CCT girls, differ based on whether they are in primary or secondary school. We again highlight that any lack of impact should be taken with caution, given the low statistical power of this analysis once sub-samples are taken – now with additional degrees of freedom lost to further controls.

Nevertheless, as above, we find, for primary schoolgirls, that the UCT decreases the likelihood of pregnancy by the end of the two years (Table 37). Its effect on marriage, however, appears less significant once these controls are included (columns 1 and 2; compare to columns 1–3 in Table 24), and the effect of the UCT on pregnancy is offset for girls who believe their best friend had sex at baseline.

Table 37: Additional variables – primary schoolgirls

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	–0.008	–0.035	0.005	0.052
	(0.012)	(0.022)	(0.017)	(0.034)
= 1 if unconditional schoolgirl	–0.022**	–0.055*	–0.014	–0.081***
	(0.011)	(0.029)	(0.018)	(0.031)
Non-zero individual earnings	0.025**	0.053***	–0.008	–0.018
	(0.010)	(0.020)	(0.015)	(0.028)
Best friend had sex	0.025**	0.039	0.046***	0.106***
	(0.012)	(0.027)	(0.016)	(0.041)
Number of observations	1,468	1,187	1,467	1,187

Note: Columns 2 and 4 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

For secondary schoolgirls, adding these controls further decreases the significance of the treatment variables on both marriage and pregnancy. Power is also lost with this adjustment (Table 38).

Table 38: Additional variables – secondary schoolgirls

	(1) Ever married R2	(2) Ever married R3	(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.046	0.084	0.056	0.100*
	(0.033)	(0.053)	(0.046)	(0.051)
= 1 if unconditional schoolgirl	–0.029	–0.086*	0.019	0.000
	(0.021)	(0.048)	(0.037)	(0.044)
Non-zero individual earnings	0.052**	0.006	0.004	–0.005
	(0.023)	(0.044)	(0.037)	(0.053)
Best friend had sex	0.004	–0.006	–0.001	0.030
	(0.027)	(0.044)	(0.037)	(0.041)
Number of observations	574	574	574	574

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.2.4 Households with more than one girl

Finally, we observe that although the outcome variables are at the individual level, participant households may have more than one girl (aged 13–22) in the study. Because treatment arm assignment is done at the EA level, all baseline schoolgirls in the same household are in the same arm. Nevertheless, we find that 26.6 percent of girls in this study live in a household with at least one other girl enrolled in the study and, alternatively, 13 percent of households in the study have more than one girl enrolled. To account for this, we re-run our preferred specification (separating primary and secondary schoolgirls), with a control variable indicating that there is another enrolled girl in the household, and with interactions between treatment and the presence of another enrolled girl.[22]

The results are largely the same for primary schoolgirls once these dummies are included, with the UCT decreasing both marriage and pregnancy (Table 39). Oddly, it appears that having another girl in the household in the UCT round also decreases the likelihood of pregnancy by the end of round 2, but not round 3.

Table 39: Baseline primary schoolgirls – marriage and pregnancy

	(1) Ever married R2	**(2)** Ever married R3	**(3)** Ever pregnant R2	**(4)** Ever pregnant R3
= 1 if conditional schoolgirl	0.006	–0.009	0.012	0.030
	(0.016)	(0.028)	(0.019)	(0.036)
= 1 if unconditional schoolgirl	–0.019	–0.062**	0.015	–0.086***
	(0.017)	(0.031)	(0.022)	(0.029)
CCT * more than one girl	–0.029	–0.052	–0.038	0.041
	(0.023)	(0.044)	(0.028)	(0.057)
UCT * more than one girl	–0.015	–0.014	–0.088***	0.013
	(0.031)	(0.048)	(0.034)	(0.051)
More than one study girl in household	–0.001	–0.021	0.002	–0.055*
	(0.017)	(0.030)	(0.018)	(0.031)
Number of observations	1,490	1,206	1,489	1,206

Note: Columns 2 and 4 exclude girls whose highest grade completed at baseline is grade 8. Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[22] We also re-run the regressions found in the paper, allowing for two levels of error clustering (EA and household). As expected, this does not change the results, since the household is nested within the EA. See Cameron and Miller (2015) for a discussion on nested levels of clustering.

In contrast to the results for primary schoolgirls, the results for secondary schoolgirls are substantially different, though it is worth reiterating that the sample size is also smaller. First, once these dummies are included, neither treatment appears to have a significant effect on pregnancy. Second, and in contrast, the effect of the UCT in decreasing marriage is even larger than in previous estimations *for households with just one girl in the study.* For households with more than one girl in the study, the effect of the UCT is erased.

Table 40: Baseline secondary schoolgirls – marriage and pregnancy

	(1) Ever married R2	**(2)** Ever married R3	**(3)** Ever pregnant R2	**(4)** Ever pregnant R3
= 1 if conditional schoolgirl	0.060	0.081	0.094	0.097
	(0.049)	(0.063)	(0.063)	(0.060)
= 1 if unconditional schoolgirl	–0.064***	–0.146***	0.012	–0.061
	(0.021)	(0.052)	(0.049)	(0.055)
CCT * more than one girl	–0.051	0.002	–0.108	0.016
	(0.052)	(0.066)	(0.079)	(0.070)
UCT * more than one girl	0.095**	0.150**	0.008	0.152*
	(0.045)	(0.071)	(0.083)	(0.090)
More than one study girl in household	–0.036*	–0.016	–0.019	–0.042
	(0.021)	(0.037)	(0.038)	(0.042)
Number of observations	586	586	586	586

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline, female-headed household at baseline, household mobile ownership at baseline, non-zero individual earnings, household school expenditures at baseline, girls' baseline assessment of likelihood of being in school the following year and completing secondary school, having received a marriage promise at baseline, wanting to get married, plan of marriage within 3 years from baseline, desired schooling, whether respondent worries about HIV/AIDS (all at baseline) and sampling strata. Errors clustered at the EA level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Altogether, these results reiterate that the benefits of the UCT in decreasing the likelihood of early pregnancies are concentrated on baseline primary girls, while decreasing early marriage for both primary and secondary schoolgirls. However, this last finding indicates that, on average, the effect in the latter group is found among girls who are the only beneficiaries of the program in the household. One possible explanation is that, although one girl's UCT receipts may protect her from the pressures of getting married/being married off, as the original authors argue in the paper, the presence of another girl who is also receiving the transfer could make marrying off the first one more feasible, since the household can get the benefits of marrying one off (including avoiding her getting "too old" for the marriage market), while still retaining a stream of transfers from the other.

5. Limitations and other analysis originally in the research plan

Besides the results above, the proposed research plan includes analyses that relied on teacher-reported school attendance and/or student test scores, and thus could not be carried out. The main missing activity in this area is an analysis of under/over-reporting of attendance by school and teacher. The original paper argues that using self-reported data leads to an overestimate of school enrollment, particularly for UCT girls and favors the use of teacher-reported data.

The pattern of the results (Baird et al. 2011, Table III) is such that, in the self-reported data, the effect of the UCT is greater than that of the CCT, and statistically significantly so in most instances, while the effect of the CCT is greater than the UCT in teacher-reported data, and the latter is often the only significant one in most instances. If the teacher-reported data is true, these results would imply not only that UCT girls over-report at a higher rate than CCT girls (as the effect of the UCT is larger than that of the CCT, using self-reported data), but also that control girls also over-report at a higher rate than CCT girls. The latter comes from the fact that the coefficient on the CCT is much larger using teacher-reported data than self-reported data, and the difference between the coefficients for the CCT treatment between the regressions using the two different data sources is significant in half the instances.

Although this is certainly possible – as CCT girls know their attendance is tracked by the research team. while girls in the other two groups do not and thus may feel more comfortable inflating their attendance rates – it would be worth conducting further analysis, including checking whether there are schools or teachers who are more likely to over-report the attendance levels of CCT girls. In particular, as the authors have both data sources, we would argue for checking whether discrepancies between the two in terms of enrollment tend to be concentrated in certain schools or teachers. As we mention in our replication proposal, the tendency for control girls in particular to inflate their attendance rates, relative to the CCT group, may be just one manifestation of a broader set of actions taken to compensate for their lack of treatment (a “John Henry effect”) that calls for further investigation.

Although proposed in the replication plan, we also could not explore the concentration of marriage and pregnancies between girls who drop out and girls who continue in school during the study period, as the analysis for this portion also relies on teacher-reported data in the paper. The original study argues that the effect of the UCT on deterring marriages and pregnancies during the period is driven by the cash transfer to girls who drop out of school (Table IX), as their counterparts in the CCT arm stop receiving the benefit by design. Dropouts, however, are defined by the teacher-reported variable, as self-reported enrollment indicates that UCT girls attend school at higher rates than CCT girls anyway, as replicated here. The replication analysis originally proposed would have used a switching regression to account for the fact that dropouts are also driven by the treatment; although we would not expect the direction of the estimates to change by employing this method, with teacher-reported data, we could perhaps obtain a more accurate measurement of the effect of the CCT on pregnancy and marriage rates for dropout and non-dropout girls separately.

Throughout the measurement and estimation analysis, we have also focused on sub-samples of girls (e.g. splitting them between primary and secondary schoolgirls as well as those with stronger or weaker school attachment). This approach has allowed us to observe heterogeneous effects, revealing stronger responses for primary schoolgirls and those with greater school attachment, particularly in terms of the likelihood of becoming pregnant by the end of the study period. That said, the lack of observed effect on secondary schoolgirls should be taken with caution, as the analysis is underpowered (well below 80%) with this sub-sample, and a true effect may have been missed.

6. Conclusion

This replication exercise relies on two sets of data from the *Cash or condition* study. The first is a clean data set (along with a do-file) available through one of the authors' websites, which allows a complete PBR. The output from the PBR is essentially the same as that presented in the original paper, and most minor discrepancies are also noted by the authors in a separate document in the same folder. The second data set, from the World Bank Microdata Library, contains all of the household survey data for the three rounds of the study (baseline, end of the first year and end of the second year). These are the data we use for both the pure replication and the measurement and estimation analysis, as they contain all of the variables collected through the surveys and have not been processed to the same extent as the clean data set. However, this World Bank data set does not include information collected through teacher surveys, school attendance ledgers or tests administered by the research team.

Not using the data collected outside the household survey substantially limits the coverage of this replication activity, as we cannot run any analysis that uses teacher-reported school enrollment or test scores as outcomes. As such, only part of the argument posed by the original paper is fully replicated here. As in the original study, we find that the UCT has a larger and more significant impact on increasing self-reported school enrollment during the study period, as well as generally decreasing marriage and/or pregnancy during the study period.

The authors of the original study argue, however, that self-reported school enrollment data is not as reliable as teacher-reported data, and use the latter along with scores from tests administered by the research team to show that CCTs actually have a larger effect on school enrollment (based on teacher reports) and learning (based on scores for an English test). Unfortunately, we cannot replicate this finding with the current data. With the available data, it would seem that UCTs generally increase (self-reported) school attendance by more than CCTs and have the additional benefit of lowering pregnancy and marriage rates among treated girls.

This difference between the study results and our replication, given the available data, actually highlights the value of using school data from more than one source (or from school ledgers and/or teacher-reported data, if indeed more reliable), as well as further investigating the reliability of teacher-reported versus self-reported data. Through our replication, we show that a study identical to *Cash and condition* that relies only on household survey data would reach the conclusion that UCTs achieve greater impact than CCTs, including in raising school attendance.

The robustness checks and analysis of heterogeneous effects we conduct generally support the results found in the original study – those we could replicate – with UCTs having a larger effect on self-reported school attendance than CCTs, as well as decreasing the likelihood of pregnancy and marriage. These two latter outcomes, however, do not seem to go perfectly hand-in-hand in response to the transfer. Among girls with a weaker link to schooling at baseline, for example, receiving the UCT decreases the likelihood that they get married during the period but not the likelihood that they become pregnant.

One possible explanation for this difference – and a point of care for policy – is that the program targeted girls who were not married at baseline. So, girls who have a history of school dropouts and/or repetition (although they were enrolled at baseline) and are thus more susceptible to decisions that could lead to permanent school abandonment – such as marriage or pregnancy – may have avoided (or their parents may have avoided their) marriage during the period, so as to not risk losing the transfer. We do not find evidence that the UCT deterred them from pregnancies, however, although the power for this sub-sample was also low. This is worth considering in policy, as a program that does not exclude married girls or that spans a longer treatment period might not have the same effect.

Separately, we find that the UCT decreased the likelihood of pregnancy among girls enrolled in primary school at baseline by 9.2 percentage points but could not detect an effect on pregnancy for secondary schoolgirls, even though the parents of this latter group received double the transfer, compared to those of primary schoolgirls (which could be used to cover school fees). In contrast, the UCT decreased the likelihood of marriage across all subgroups of girls.

Altogether, the impact of the UCT in averting pregnancies was concentrated on girls with greater school attachment and girls in primary school (likely younger girls), while it discouraged marriages for girls with weaker attachment and girls in secondary school. In this sense, with limited resources, a UCT targeting one of the former groups may be most effective for decreasing teen pregnancies and marriages, although including the latter group can also decrease marriage rates for them. That said, the fact that the study included only unmarried girls, by design, may pose a limitation to the translatability of these results for policy, as girls might have hesitated to marry (or their parents to marry them off) if they were concerned that it would lead to their removal from the program. An unconditional transfer program that targets all girls in primary school (or any other group discussed above) might not have the same result if it were not restricted to unmarried girls.

Also, the UCT's ability to discourage marriages among older girls appears to be concentrated in households where these girls are the only beneficiaries of the program. In this sense, concentrating the treatment on younger girls could even be preferable in the face of limited resources.

An additional program design consideration pertains to the transfer amounts. Although distribution of the randomized amounts was the same between UCT and CCT arms, households in the UCT arm with girls in secondary school at baseline received an additional transfer, equivalent to the average cost of school fees. This was done for comparability with the CCT arm, under which the research team covered the secondary school fees directly, but represents a much larger transfer, relative to the control group. On average, this additional transfer, doubled the amount given to parents in this arm, and, nevertheless, decreased pregnancies but not marriage rates for girls in this group. Taking into account that the total transfer was around US$17 per household (parent and girl) with secondary schoolgirls, rather than US$10 (the mean amount for primary schoolgirls), increases the cost-benefit ratio of including secondary schoolgirls in a targeted program.

Although exploratory, we also call attention to the fact that in some of our regressions for the likelihood of pregnancy among older girls and those with weaker attachment to school, the coefficient of the CCT was positive, large (ranging from 8 to 10 percentage points), and significant at least at the 90 percent level. The samples from which this impact was found are small, but this result may serve as a warning for implementing CCTs as policy and might call for a check that they are not increasing pregnancy rates among older girls.

Another policy-relevant finding is the spillover effect on non-treated girls, along with the absence of an additional effect on treated girls from the variation in the concentration of transfers within EAs. That is, the effect of the UCT in decreasing marriages and pregnancies during the study period was not statistically different for girls receiving the transfer directly and those in the same EA who did not receive it. This indicates that the geographic concentration of the program might matter for these outcomes, at least over the span of two years. Ignoring concerns about fairness in selection and distribution, a transfer program that provides UCTs to just one-third of eligible schoolgirls could have effects on decreasing pregnancy and marriages across all eligible girls that are similar to one that actually makes transfers to all eligible girls. There is no additional effect on treated girls from having more or fewer girls nearby who also receive the transfer. That said, this spillover effect could dissipate over time, as non-treated girls decrease their expectation that they could soon be invited to enroll in the program.

Lastly, we point to the robustness of the original results on marriage and pregnancy to the inclusion of other control variables. The original study includes several baseline control variables in its regressions, such as age, asset index and highest grade achieved to that point. In this replication, we explore some other variables that could have influenced the outcomes of interest and were unbalanced across treatment arms, such as a girl's individual earnings, whether her best friend had sex in the year prior, and her aspiration for marriage and the amount of education acquired before that. When we re-run the analyses including these variables, we find that the coefficients for the treatment variables are qualitatively the same as in the original study. Separately, we also find that girls who have some earnings at baseline are more likely to be married by the end of the study. And girls whose best friend had sex – and so are presumably more likely to find sex outside marriage or at a younger age normal – are also more likely to have ever been pregnant by the end of the two years.

In this replication, given our data limitations, we have focused on self-reported school enrollment and the likelihood of pregnancies and marriage. With respect to these outcomes, our results indicate that a UCT is more effective than a CCT, increasing school enrollment (although the reliability of this measurement may be questionable) and decreasing the likelihood that a beneficiary girl has ever gotten pregnant or married during the two-year study period. This points to clear benefits from an unconditional transfer, with girls in primary school (or with a stronger school attachment) experiencing impacts with respect to both outcomes, with those in secondary school (or with a weaker school attachment) only avoiding marriage (at best) in response.

The finding of the original study on the positive average impacts of the UCT on marriage and pregnancies is important for the literature on unconditional transfers and on comparing these with CCTs, and highlights the fact that UCTs may have an impact on different areas than CCTs, as they offer support to those who might not be able or willing to comply with a conditionality. This replication not only confirms the positive impact of the UCT on decreasing marriage and pregnancies – in contrast to the CCT, which has no effect on these two outcomes – but also pinpoints primary schoolgirls and those with stronger school attachment as the greater beneficiaries.

Appendix: PBR output and comparison, by table

This appendix presents a replication of each table in the paper.[23] Whenever there are differences between the results in the paper and the PBR output, both are shown. The results in the paper are in black; those derived from the PBR, if different, are set in **bold**.

Table I

For Table I, the results largely match each other and are qualitatively identical. A slight difference between the estimates presented in the paper and the PBR output can be found in column 6. This is likely caused by a dropped observation in the PBR.

Table A1: PBR of Table I

	Three round panel, education exists[24]						
	(1) Round 3	(2) Panel	(3) Education test	(4) School survey R2	(5) School survey R3	(6) Has legible ledger	
= 1 if conditional schoolgirl	0.020	0.021	0.029*	0.033	–0.000	0.116*	**0.115***
	(0.015)	(0.030)	(0.016)	(0.024)	(0.027)	(0.064)	
= 1 if unconditional schoolgirl	0.021	0.030	0.035*	–0.029	0.014	0.061	**0.060**
	(0.019)	(0.024)	(0.020)	(0.053)	(0.028)	(0.077)	
Control group mean	0.946	0.893	0.929	0.890	0.935	0.378	**0.379**
	(0.009)	(0.011)	(0.010)	(0.015)	(0.018)	(0.039)	
Number of observations	2,284	2,284	2,284	2,284	983	821	**820**
Prob > F (conditional = unconditional)	0.965	0.797	0.801	0.246	0.627	0.513	

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table II

Similarly, the results for Table II achieved through the PBR are qualitatively the same as those presented in the paper, although there is one minor difference for one variable.

We also note that the replication do-file made available by the authors does not provide the code for getting the means and standard errors for the household and individual transfers. That said, this is fairly straightforward and could be done manually by the replication team, following the same format used for getting these estimates for the other variables in the table.[25] The cells requiring this manual replication are presented in

[23] The paper also includes nine tables in a supplementary online appendix, but there does not appear to be a PBR code available for these tables (or the research team could not find it).
[24] The authors' memo shows the same differences.
[25] In particular, we first de-normalize the variables:
gen hh_amount_edit=hh_amount_USD+4
gen individual_amount_edit=individual_amount_USD+1

italics, although they match the results in the paper. Interestingly, in the memo, the authors present a different standard error for these computed cells from what is presented in the paper and from the PBR output.

Table A2: PBR of Table II

	Baseline means and balance[26]			
	(1) Control group mean	(2) = 1 if conditional schoolgirl	(3) = 1 if unconditional schoolgirl	(4) p-value (conditional-unconditional)
Panel A: Household-level variables				
Household size	6.432 (2.257)	6.384 (2.146)	6.662 (2.075)	0.202
asset_index	0.581 (2.562)	0.984 (2.740)	1.221 (2.447)	0.623
female_headed	0.343 (0.475)	0.252** (0.434)	0.245** (0.431)	0.899
mobile	0.616 (0.487)	0.583 (0.494)	0.605 (0.490)	0.799
hh_amount	n/a	*6.991* *(2.319)*	*6.829* *(2.101)*	0.822
Panel B: Individual-level variables				
age_R1	15.252 (1.903)	14.952* (1.827)	15.424 (1.923)	0.007***
highest_grade	7.478 **7.479** (1.634) **(1.633)**	7.246 (1.598)	7.896** (1.604)	0.004***
mother_alive	0.842 (0.365)	0.802 (0.399)	0.836 (0.371)	0.360
father_alive	0.705 (0.456)	0.714 (0.453)	0.759 (0.428)	0.288
never_had_sex (round 1)	0.797 (0.402)	0.797 (0.403)	0.775 (0.419)	0.582
ever pregnant (round 1)	0.023 (0.149)	0.030 (0.171)	0.031 (0.173)	0.973
individual_amount	n/a	*3.090* *(1.431)*	*3.033* *(1.451)*	0.606
Observations	1,356 **1356–1358**	470 **468–470**	261 **259–261**	

The number of observations is not the same across all variables, as implied by the last row in Table II in the paper, though noted otherwise in the table's footnote. The range of observations are presented **bolded** here.

Then, we use the following code to get each of the means and standard deviations:

```
foreach x in hh_amount_edit individual_amount_edit {
        sum `x' [aw=wgt] if T2a==1 & round==1 & sample_SG==1 & panel==1;
        sum `x' [aw=wgt] if T2b==1 & round==1 & sample_SG==1 & panel==1;
}
```

[26] The authors' memo shows different standard errors than those we computed (and those presented in the paper) and do not show the range in the number of observations presented here.

Table III

There are a few minor differences in Panel A for Table III, likely stemming from dropped observations from the paper to the replication file. Qualitatively, the results are essentially similar, with a minor shift in column 5 from the difference between conditional and unconditional schoolgirls being statistically insignificant at conventional levels to significant at the 90 percent level. There are no differences in Panel B between the results presented in the paper and those coming from the PBR exercise.

Table A3: PBR of Table III

Panel A

	Enrollment – self-reported[27]							
	(1) 2008 Term 1	(2) 2008 Term 2	(3) 2008 Term 3	(4) 2009 Term 1	(5) 2009 Term 2	(6) 2009 Term 3	(7) TOTAL number of terms	(8) 2010 Term 1
= 1 if conditional schoolgirl	0.007	0.019*	0.041**	0.049*** **0.048*** **	0.056***	0.061***	0.233*** **0.232*** **	0.005
	(0.011)	(0.011)	(0.017)	(0.017)	(0.018)	(0.019)	(0.070)	(0.025)
= 1 if unconditional schoolgirl	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 of 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 **2,085**	2,087 **2,085**	2,087 **2,086**	2,086 **2,084**	2,086
Prob > F (conditional = unconditional)	0.006	0.012	0.460	0.299 **0.293**	0.102 **0.100**	0.098 **0.095**	0.038 **0.037**	0.028

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

[27] The authors' memo shows the same differences.

Panel B

	Enrollment – school survey							
	(1) 2008 Term 1	(2) 2008 Term 2	(3) 2008 Term 3	(4) 2009 Term 1	(5) 2009 Term 2	(6) 2009 Term 3	(7) TOTAL number of terms	(8) 2010 Term 1
= 1 if conditional schoolgirl	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)
= 1 if unconditional schoolgirl	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 of 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

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table IV

As noted in the paper, the results presented in Table IV are actually from a different paper by a subset of the co-authors, which looks at the reliability of self-reported versus school-reported attendance rates (Baird and Özler 2012). As such, the do-file for the replication of the present *Cash and condition* paper does not contain the code for Table IV. However, the replication team was able to create a version of this table by using the code available in the replication do-file for the other paper.[28]

Table A4: PBR of Table IV

	Reporting bias on enrollment	
	(1) Core respondents over-reporting	(2) Teachers over-reporting
= 1 if conditional schoolgirl	–0.093**	–0.021
	(0.052)	(0.035)
= 1 if unconditional schoolgirl	–0.001	–0.014
	(0.058)	(0.038)
Mean in control group	325	325
Number of observations	0.170	0.052
Prob > F (conditional = unconditional)	0.018	0.794

[28] The replication file for Baird and Özler (2012) can be found on Berk Özler's website: https://drive.google.com/open?id=0B274-JLBCKcdcWxwbjIyX0E0b00. The results in Tables IV in Baird and colleagues (2011) and in Baird and Özler (2012) differ slightly, because the former (replicated here) do not include any controls – unlike its other regression tables – whereas the latter includes the standard controls used throughout the paper (e.g. age and highest grade achieved at baseline).

Tables V and VI

All the numbers presented in Tables V and VI in the paper match the PBR output.

Table A5: PBR of Table V

	Attendance – school ledgers				
	(1) Ledger 2009-T1	(2) Ledger 2009-T2	(3) Ledger 2009-T3	(4) Overall 2009	(5) Ledger 2010-T1
= 1 if conditional schoolgirl	0.139***	0.014	0.169**	0.080**	0.092**
	(0.045)	(0.033)	(0.085)	(0.035)	(0.041)
= 1 if unconditional schoolgirl	0.063	0.038	0.118	0.058	-0.038
	(0.056)	(0.033)	(0.102)	(0.037)	(0.053)
Mean of 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

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A6: PBR of Table VI

	Educational tests			
	(1) English Score (standardized)	(2) TIMMS Score (standardized)	(3) Math Malawi score (standardized)	(4) Cognitive score (standardized)
= 1 if conditional schoolgirl	0.140***	0.120*	0.086	0.174***
	(0.054)	(0.067)	(0.057)	(0.048)
= 1 if unconditional schoolgirl	–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	0.7556

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table VII

In Table VII, there are slight differences in the estimates relating to whether a girl was ever married by round 3.[29] This is likely due to a smaller number of observations used in the paper but does not render the results meaningfully different from each other.

Table A7: PBR of Table VII

	Marriage and pregnancy[30]				
	(1) Ever married R2	(2) Ever married R3		(3) Ever pregnant R2	(4) Ever pregnant R3
= 1 if conditional schoolgirl	0.007	–0.012	**-0.006**	0.013	0.029
	(0.012)	(0.024)		(0.014)	(0.027)
= 1 if unconditional schoolgirl	–0.026**	–0.079***	**-0.076*** **	–0.009	–0.067***
	(0.012)	(0.022)		(0.017)	(0.024)
Mean of control group	0.043	0.180	**0.176**	0.089	0.247
Number of observations	2,087	2,084	**2,086**	2,086	2,087
Prob > F (conditional = unconditional)	0.024	0.025	**0.019**	0.265	0.003

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table VIII

Table VIII is a descriptive table showing the marriage rates by enrollment status and treatment group. There are discrepancies between the results in the paper and the PBR output for almost every cell. The fact that the row percentages are different for all arms (control, conditional treatment and unconditional treatment) indicates that the number of observations is probably different between the paper and the PBR file for each of these arms. Nevertheless, the results are essentially the same, showing that marriage rates are much higher among non-enrolled than enrolled girls but also among non-enrolled girls in the control and conditional arms than in the unconditional arm.

That said, in all cells except one, the share of girls who are married in the PBR output is smaller than that presented in the paper, and, in total, whereas the paper shows that 17.2 percent of girls in these three arms were ever married by the beginning of 2010, the PBR results finds only 14.4 percent of girls are married.

---

29 The discrepancies relating to the "ever married" outcome are acknowledged in the authors' memo, included with the replication files. They explain that these are due to "additional cleaning of the data post-publication" and note that "[a]ll results remain qualitatively the same."
30 The authors' memo shows the same differences.

Table A8: PBR of Table VIII

Marital status by enrollment and group[31]	(1) Enrolled		(2) Not enrolled		(3) Total	
Control, %	1.7	**0.4**	46.9	**42.4**	19.9	**18.2**
(row %)	(59.8)	**(57.5)**	(40.2)	**(42.5)**	(100.0)	
Conditional treatment, %	0.5		50.8	**48.7**	16.0	**16.1**
(row %)	(69.2)	**(67.6)**	(30.8)	**(32.4)**	(100.0)	
Unconditional treatment, %	0.3		25.2	**23.0**	10.1	**9.4**
(row %)	(60.5)	**(59.9)**	(39.5)	**(40.1)**	(100.0)	
Total, %	1.1	**0.4**	44.2	**37.7**	17.2	**14.4**
(row %)	(62.7)	**(62.4)**	(37.3)	**(37.6)**	(100.0)	

Table IX

Most of the push-button output for Table IX does not match that of the paper. Although the findings are generally the same, these differences do lead to some changes in statistical significance at conventional levels. In particular, the second column uses "ever married" by the beginning of 2010 as the dependent variable. In the paper, the difference between the effect of conditional and unconditional treatment is rejected at conventional levels, while the push-button results show them to be different at the 90 percent confidence level. The opposite is true for the first column, which uses an indicator of teacher-reported enrollment as the dependent variable; it switches from barely significant to not significant at conventional levels.

Although the discrepancies in the estimates are likely driven, in part, by the different number of observations in the paper and the PBR data set for columns 1, 2 and 4, the discrepancies in the third column are slightly more puzzling, as the number of observations is the same between the two files.

Nevertheless, the main messages used from this table in the paper – that the UCT affects marriage rates but not school dropouts, and that the UCT is associated with lower marriage rates among non-enrolled girls but not among enrolled girls – hold just the same in the PBR output.

[31] The authors' memo shows the same differences.

Table A9: PBR of Table IX

	Teacher-reported school enrollment and marital status[32]							
	(1)		(2)		(3)		(4)	
	In school		Ever married		Ever married		Ever married	
	All		All		Enrolled		Not enrolled	
= 1 if conditional schoolgirl	0.058*		–0.026	**–0.015**	–0.012	**0.002**	0.033	**0.026**
	(0.034)	**(0.033)**	(0.037)	**(0.036)**	(0.015)	**(0.005)**	(0.097)	**(0.088)**
= 1 if unconditional schoolgirl	–0.000	**0.001**	–0.088***	**–0.081*** **	–0.011	**–0.002**	–0.159**	**–0.155*** *
	(0.036)		(0.030)		(0.010)	**(0.005)**	(0.067)	**(0.066)**
Mean of control group	0.598	**0.596**	0.199	**0.191**	0.017	**0.004**		**0.469**
Number of observations	844	**847**	844	**851**	490		354	**356**
Prob > F (conditional = unconditional)	0.099	**0.108**	0.106	**0.084**	0.857	**0.403**	0.088	**0.078**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

The authors' memo shows the same discrepancies presented here for the first regression in Table IX but not for the other three. However, there is a note below the table in their memo indicating that it needs to be updated as more of the cells do not match.

Table X

The results for three of the regressions presented in Table X are identical between the paper and the PBR output, but those for regression 3, which has "ever married" by round 3 as the dependent variable, are slightly different. Again, this is most likely because the paper uses two fewer observations than the PBR file (as in Table VII, above). Nevertheless, the results are qualitatively similar, although the difference between the CCT interacted with age and the UCT interacted with age increases, as the coefficient for the interaction term between UCT and being above 15 years old becomes more negative and significant. This is consistent with the authors' argument that UCTs are a more effective tool for reducing marriage rates, particularly among older girls.

[32] The authors' memo shows the same differences for the first column but not for the subsequent columns. They have a note in their memo stating that their table needs to be updated.

Table A10: PBR of Table X

	Age heterogeneity[33]				
	(1)	(2)	(3)		(4)
	Enrollment	English	Ever married		Ever pregnant
= 1 if conditional schoolgirl	0.467***	0.141*	–0.023	**–0.017**	–0.008
	(0.159)	(0.073)	(0.017)		(0.028)
= 1 if unconditional schoolgirl	0.257	–0.116	–0.051**	**–0.046****	–0.059***
	(0.157)	(0.102)	(0.020)	**(0.021)**	(0.020)
Conditional treatment * above 15 years old	0.290	0.017	0.037	**0.036**	0.104*
	(0.291)	(0.089)	(0.056)	**(0.055)**	(0.054)
Unconditional treatment * above 15 years old	0.103	0.245**	–0.067	**–0.071***	–0.032
	(0.255)	(0.110)	(0.042)		(0.046)
= 1 if above 15 years old	–0.786***	–0.546***	0.122***	**0.126*****	0.176***
	(0.244)	(0.058)	(0.026)		(0.027)
Number of observations	852	2,057	2,084	**2,086**	2,087
Prob > F (conditional = unconditional)	0.095	0.031	0.188	**0.184**	0.067
Prob > F (conditional * older = unconditional * older)	0.364	0.059	0.097	**0.087**	0.027

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table XI

Although they are qualitatively similar, almost none of the estimates in Table XI match between those presented in the paper and the PBR output. This is surprising, given that in three of the four regressions in this table, the number of observations used are the same in the two sources. Nevertheless, there are no major qualitative differences, although a few coefficients differ in significance (although less so in magnitude). Most notably, perhaps, in shifting from the paper to the PBR, the coefficient on the minimum transfer for the UCT increases from 9 to 10 percentage points and moves up (from p < 10% to p < 5%) on the regression for "ever pregnant." That is the only statistically significant effect documented in this table on whether a girl has ever been pregnant.

[33] The authors' memo shows the same differences.

Table A11: PBR of Table XI

	Transfer amounts[34]							
	(1)		(2)		(3)		(4)	
	Total terms enrolled		English test score		Ever married		Ever pregnant	
Conditional treatment, individual amount	0.024	**0.040**	–0.032	**–0.023**	–0.002	**-0.003**	0.006	**0.004**
	(0.051)	**(0.052)**	(0.029)	**(0.026)**	(0.008)		(0.012)	
Unconditional treatment, individual amount	–0.048	**–0.054**	–0.019	**–0.020**	–0.016	**-0.014**	0.013	
	(0.064)	**(0.067)**	(0.038)	**(0.039)**	(0.011)		0.013	**(0.014)**
Conditional treatment, household amount	–0.027	**–0.039**	–0.000	**–0.004**	0.001	**0.002**	0.005	**0.006**
		(0.035)	(0.016)	**(0.017)**	(0.007)		(0.010)	
Unconditional treatment, household amount	0.081***	**0.074****	–0.058**	**–0.059****	–0.017**	**-0.016*****	–0.002	
	(0.031)	**(0.034)**	(0.029)	**(0.027)**	(0.007)	**(0.006)**	(0.009)	**(0.008)**
Conditional treatment, minimum transfer	0.572***	**0.604*****	0.202*	**0.209***	–0.011	**-0.004**	0.001	**–0.000**
	(0.213)	**(0.217)**	(0.118)	**(0.119)**	(0.044)	**(0.043)**	(0.052)	**(0.051)**
Unconditional treatment, minimum transfer	0.094	**0.192**	0.175	**0.205**	0.001	**-0.005**	–0.089*	**–0.097****
	(0.167)	**(0.179)**	(0.132)	**(0.141)**	(0.040)	**(0.041)**	(0.050)	**(0.049)**
Number of observations	852		2,057		2,084	**2,086**	2,087	
Prob > F (conditional = unconditional), individual	0.390	**0.275**	0.788	**0.949**	0.300	**0.419**	0.702	**0.613**
Prob > F (conditional = unconditional), household	0.025		0.082	**0.088**	0.069	**0.060**	0.614	**0.504**
Prob > F (conditional = unconditional), minimum	0.046	**0.086**	0.877	**0.982**	0.834	**0.983**	0.203	**0.169**

Note: Includes controls for highest grade at baseline, asset index at baseline, never had sex at baseline, age at baseline and sampling strata. Errors clustered at the EA level. * p < 0.10, ** p < 0.05, *** p < 0.01.

[34] The authors' memo only presents the differences for column 3, but theirs does not match those achieved through this PBR exercise. It seems that the authors may be aware of these differences, as they have a note in the memo stating, "the numbers don't match" for column 3.

References

Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2009. *Research Proposal: Unpacking the Impacts of a Randomized CCT program in Sub-Saharan Africa*. Available at: http://microdata.worldbank.org/index.php/catalog/2339/download/34881 [Accessed 6 December 2018].

Baird, S, Ferreira, FHG, Özler, B and Woolcock, M, 2014. Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1), pp.1–43.

Baird, S, Garfein, RS, McIntosh, CT and Özler, B, 2012. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomized trial. *The Lancet*, 379, pp.1320–1329.

Baird, S, McIntosh, C and Özler, B, 2011. Cash or condition? Evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4), pp.1,709–1,753.

Baird, S and Özler, B, 2012. Examining the reliability of self-reported data on school participation. *Journal of Development Economics*, 98, pp.89–93.

Bastagli, F, Hagen-Zanker, J Harman, L, Barca, V, Sturge, G, Schmidt, T and Pallerano, L, 2016. *Cash transfers: what does the evidence say? A rigorous review of programme impact and of the role of design and implementation features*. London: Oxford Development Institute.

Brown, AN, Cameron, DB and Wood, BDK, 2014. Quality evidence for policymaking: I'll believe it when I see the replication. *Journal of Development Effectiveness*, 6(3), pp.215–235.

Cameron, AC and Miller, DL, 2015. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), pp.317–372.

Little, RJA and Rubin, DB, 2002. *Statistical Analysis with Missing Data*. Hoboken, NJ: Wiley.

Reimão, ME, 2017. *Technical Proposal: Replication of "Cash or condition? Evidence from a cash transfer experiment."* 3ie Replication Window 4: Financial Services for the Poor. Available at: <http://www.3ieimpact.org/media/filer_public/2017/07/26/reimao-replication-plan.pdf> [Accessed 6 December 2018].

World Bank, 2010. *The Education System in Malawi: Country Status Report*. Washington, DC: World Bank, UNESCO, Pole de Dakar, Education for All and GTZ.

Other publications in the 3ie Replication Paper Series

The following papers are available from http://3ieimpact.org/evidence-hub/publications/replication-papers

Impact of unconditional cash transfers: a replication study of the short-term effects in Kenya. 3ie Replication Paper 20. Wang, H, Qui, Fang, Q and Luo, J, 2018.

Mobile money and its impact on improving living conditions in Niger: a replication study. 3ie Replication Paper 19. Meneses, JP, Ventura, E, Elorreaga, O, Huaroto, C, Aguilar, G, and Beteta, E, 2018.

Savings revisited: a replication study of a savings intervention in Malawi, 3ie Replication Paper 18. Stage, J and Thangavelu, T 2018.

Thou shalt be given…but how? A replication study of a randomized experiment on food assistance in northern Ecuador, 3ie Replication Paper 17. Lhachimi, SK and Seuring, T, 2018.

Preventing HIV and HSV-2 through improving knowledge and attitudes: a replication study of a multicomponent intervention in Zimbabwe. 3ie Replication Paper 16. Hein, NA, Bagenda, DS and Yu, F, 2018.

PEPFAR and adult mortality: a replication study of HIV development assistance effects in Sub-Saharan African countries. 3ie Replication Paper 15. Hein, NA, Bagenda, DS and Luo, J, 2018.

When to start ART? A replication study of timing of antiretroviral therapy for HIV-1-associated Tuberculosis. 3ie Replication Paper 14. Djimeu, EW, 2018.

STRETCHing HIV treatment: a replication study of task shifting in South Africa. 3ie Replication Paper 13. Chen, B and Alam, M, 2017.

Cash transfers and HIV/HSV-2 prevalence: a replication of a cluster randomized trial in Malawi. 3ie Replication Paper 12. Smith, LM, Hein, NA and Bagenda, DS, 2017.

Power to the people?: a replication study of a community-based monitoring programme in Uganda, 3ie Replication Paper 11. Donato, K and Garcia Mosqueira, A (2016)

Fighting corruption does improve schooling: a replication study of a newspaper campaign in Uganda, 3ie Replication Paper 10. Kuecken, M, and Valfort, MA (2016)

The effects of land titling on the urban poor: a replication of property rights, 3ie Replication Paper 9. Cameron, Drew B, Whitney, Edward M and Winters, Paul C (2015)

Male circumcision and HIV acquisition reinvestigating the evidence from young men in Kisumu, Kenya, 3ie Replication Paper 8. Djimeu, EW, Korte, JE and Calvo, FA (2015)

Walking on solid ground: a replication study on Piso Firme's impact, 3ie Replication Paper 7. Basurto, MP, Burga, R, Toro, JLF and Huaroto, C (2015)

The impact of India's JSY conditional cash transfer programme: A replication study, 3ie Replication Paper 6. Carvalho, N and Rokicki, S (2015)

Recalling extra data: A replication study of finding missing markets, 3ie Replication Paper 5. Wood, BDK and Dong, M (2015)

The long and short of returns to public investments in fifteen Ethiopian villages, 3ie Replication Paper 4. Bowser, WH (2015)

Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 2: Alternative analyses, 3ie Replication Paper 3, part 2. Aiken, AM, Davey, C, Hayes, RJ and Hargreaves, JR (2014)

Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 1: A pure replication, 3ie Replication Paper 3, part 1. Aiken, AM, Davey, C, Hargreaves, JR and Hayes, RJ (2014)

TV, female empowerment and demographic change in rural India, 3ie Replication Paper 2. Iversen, V and Palmer-Jones, R (2014)

Quality evidence for policymaking: I'll believe it when I see the replication, 3ie Replication Paper 1. Brown, AN, Cameron, DB and Wood, BDK (2014)

Replication Paper Series

International Initiative for Impact Evaluation
1029 Vermont Avenue, NW
Suite 1000
Washington, DC 20005
USA

replication@3ieimpact.org
Tel: +1 202 629 3939

www.3ieimpact.org