Supplementary Online Appendix

This Version: September 18, 2018

Contents

1 Supplementary Online Appendix 2

2 Review of Literature on Conditional Incentives and Empowerment Interventions 2

3 Survey Questionnaire and Variable Construction 9

4 Merging Accuracy 19

5 Randomization 20

6 Take-up Adjusted for Crossovers 21

7 Cost-Benefit and Cost-Effectiveness Analyses 22

7.1 Interventions Included in Analyses 22

7.2 Methodology 24

1 Supplementary Online Appendix

2 Review of Literature on Conditional Incentives and Empowerment Interventions

Search Process and Criteria for Inclusion in Literature Review

We searched first for "literature reviews" and "systematic reviews" of the literature on child marriage and early marriage. We confined our search to quantitative studies that attempted to evaluate the impact of programs, policies, or interventions that had child marriage or teenage childbearing as an outcome. We also searched the terms "child marriage" and "early marriage". We used the snowball method, reading the studies rated medium or high-quality in the literature reviews and the studies mentioned in the papers listed in the literature review and found from the original search terms. We searched the database of randomized control trials on the Abdul Latif Jameel Poverty Action Lab website, the International Initiative for Impact Evaluation, the Social Science Registry of randomized trials, and the Impact Evaluation Database at the World Bank for studies that had child marriage or teenage childbearing as an outcome.

We found five reviews of the literature on child marriage of which one was published in a peer reviewed journal (Lee-Rife et al., 2012; Malhotra et al., 2011; Gupta et al., 2008; Mensch and Lloyd, 2005; Parsons and McCleary-Sills, 2005) (Note that we did not specifically search for literature reviews or studies on teenage pregnancy unassociated with early marriage. The contexts where teenage pregnancy mainly occurs outside early marriage are quite different from our context and the drivers of teenage pregnancy outside marriage may be somewhat different from the drivers of child marriage. Our search terms did lead us to read and review some studies where teenage pregnancy was the main outcome. The conclusions we drew from these seem consistent with the wider literature on teenage pregnancy). There was a large overlap in the studies contained in three of the reviews on early/child marriage (Lee-Rife et al., 2012; Malhotra et al., 2011; Gupta et al., 2008), a fourth mainly reviewed (and added to) work on the correlates of child marriage (Mensch and Lloyd, 2005), and a fifth only covered studies funded by the World Bank (Parsons and McCleary-Sills, 2005). Many of the studies covered in these reviews and found in our search process were program evaluation reports which had not gone through peer review and contain little detail on methodology or results. For example, results were often provided in charts with no details on variance, standard errors, confidence intervals, or significance.

Another common issue in many studies was the few number of clusters analyzed. When program assignment is at the cluster level (e.g. community or district), analysis of program impacts must take into account cluster level shocks. This can be done by treating each cluster as one observation or by treating each individual as a separate observation but clustering standard errors at the level of the cluster. Several of the studies summarized in previous literature reviews appear to ignore in the analysis that assignment was at the cluster level. In particular, several studies have one or two treatment units and one or two comparison units. It is not possible to adjust for cluster level shocks with this limited number of clusters.

In our search we found three studies that evaluated a conditional incentive program where families received support if girls avoided child marriage even if they had dropped out of school. While two of these did not have child marriage as an outcome, and there were some quality concerns with two of the studies, we nevertheless included them in our review given the similarity to the program we evaluated.

Because of the very low quality of even those studies ranked "medium" in previous literature reviews, we did not formally review studies ranked as low-quality in these reviews.[1] The wider search process did, however, reveal several high-quality studies not previously included in other literature reviews on child marriage.

Studies were rated as low-quality if they fit any of the following criteria:

i. there were too few clusters to estimate cluster errors. Several studies had only one or two treatment or comparison areas making it impossible to compare outcomes across treatment and comparison areas;

ii. there was insufficient information on the methodology to judge its quality and/or there was no information on the statistical significance of the results;

iii. there was potential for selection bias between those who participated in the program and those that did not and there was no credible methodology of addressing this selection bias. For example, Krishnan et al. (2014) evaluate the effect of a state-wide program in Haryana by comparing marriage age of girls in Haryana with daughters-in-law living in Haryana but originally from other states. However, girls who marry into another state are likely to be different on many dimensions to those who marry within their state.

Studies were ranked medium-quality if they:

i. provided sufficient information to judge the quality of the study, the risk of selection bias, and level of statistical significance or reported confidence intervals;

ii. used appropriate methodologies to address concerns about selection bias;

iii. were unable to address all concerns about selection bias. Note that most studies ranked medium-quality were explicit that some concerns about selection bias remained.

Studies were ranked high-quality if they:

[1] We did, however, review low-quality studies that are very similar to one of our treatment arms.

i. provided sufficient information to judge the quality of the study, the risk of selection bias, and level of statistical significance or reported confidence intervals;

ii. satisfactorily addressed selection bias concerns. Many (though not all) high-quality studies used random assignment with large numbers of units or individuals randomized.

We did not use publication in peer reviewed journals as a criterion for quality as some of the studies published in peer review journals had serious methodological flaws while there were several studies with sufficient detail to judge quality but that were not published in peer reviewed journals. The latter may reflect the long lag in economics between the completion of full, academic style, working papers and final publication in a journal.

Excluding low-quality studies leaves us with nine high or medium quality studies. In addition, we consider five low-quality studies. Six studies are randomized control trials, six studies are in Africa, one is in Latin America, and seven are in South Asia (assessing three interventions). Table S1 presents all evaluated studies.

Before summarizing results, it is worth noting that there are cultural differences across regions which are important to keep in mind when comparing results. In particular, teenage childbearing without formal marriage is more common in some parts of Africa than it is in South Asia but usually means that girls cannot stay in school. As child marriage is often associated with teenage childbearing there are important parallels between these two outcomes. Another important relevant difference to keep in mind is that in some countries marriage timing is mainly determined by parents of girls whereas in other countries girls themselves may have more say over cohabitation, marriage, or sexual onset.

Evidence on the Impact of Unconditional Incentives

There is very little evidence on the impact of unconditional incentives on child marriage or teenage childbearing and existing evidence is mostly from Africa. One high-quality study evaluates an RCT in Malawi in which girls and their families received monthly stipends over the course of two school years (Baird et al., 2011). The incentives significantly reduced teenage childbearing and marriage rates. Another low-quality study evaluates an RCT to assess the impact of unconditional cash incentives to orphans in Kenya (Handa et al., 2015). While the program reduced the likelihood of pregnancy, there was no significant impact on early marriage. However, respondents were selected differently within treatment and control arms, leading to systematic differences in characteristics between the two groups. There was substantially higher attrition in the control group than in the treatment group (30% vs. 17%), and the study had overall relatively few clusters (14 in treatment and 14 in control).

Evidence on the Impact of Incentives Conditional on Education

Cash and noncash incentives conditional on girls staying in school and/or maintaining good grades appear to be effective in encouraging girls to stay in school but there is mixed evidence on whether this leads to reductions in early marriage, teen pregnancy, or cohabitation during teenage years (note that we did not explicitly search for studies having teen pregnancy or cohabitation as an outcome). In total, seven studies evaluate the impact of such programs and show significant positive effects on schooling (Baird et al., 2011; Duflo et al., 2015; Alam et al., 2011; Hong and Sarr, 2012; Hahn et al., 2015; Angrist et al., 2006; Heath and Mobarak, 2015). Five studies have negative significant effects on marriage, teen pregnancy, or cohabitation, (Duflo et al., 2015; Alam et al., 2011; Hong and Sarr, 2012; Hahn et al., 2015; Angrist et al., 2006), and one has effects that are insignificantly different from zero (Heath and Mobarak, 2015). If we focus on the three high-quality studies, two have significant reductions in pregnancy or cohabitation although neither has formal marriage as an outcome, (Duflo et al., 2015; Angrist et al., 2006) and one has an effect insignificantly different from zero (Baird et al., 2011). It is worth noting that neither of the high-quality studies linking conditionality on education to marriage are in South Asia or even in countries where parents are the main deciders of marriage age. One is a study of free school uniforms in Kenya (Duflo et al., 2015) and one assesses the impact on teen pregnancy and marriage rates of incentives conditional on school attendance in Malawi (Baird et al., 2011). One evaluates a program in which vouchers to attend private school (which are valuable) are conditional on girls maintaining higher grades in Colombia (Angrist et al., 2006). It is not possible to separate the effect of attending a private school on cohabitation (the outcome in this study) from the effect of the conditionality on grades. The study does, nevertheless, suggest a causal link between education and marriage/cohabitation. The context however, is very different from that in South Asia. The four low and medium-quality studies examining conditional incentives and their link to child marriage are more closely tied to the context of our study. All are in South Asia, have marriage as an outcome, and evaluate Female Secondary School Stipends (FSSP) in Pakistan or Bangladesh where parents have considerable control over marriage. None, however, completely address concerns of selection bias and they find mixed results. Alam et al. (2011) compare marriage trends for girls in districts in Pakistan where the FSSP was introduced with girls in districts where it was not introduced. They find large effects on education and marriage but these are not very robust to different specifications and results are insignificant in the purest intention-to-treat specification. Program districts were chosen because education of girls was particularly low in these districts. This raises the concern that outcomes for poor performing districts would have risen faster than better performing districts in a process of catchup even without the program. In addition, there is very strong geographic clustering of treatment and control districts with virtually all treatment regions in the South of Punjab and virtually all control districts in the North, meaning that identification relies only on the north south divide, reinforcing concerns that the result could be driven by differential trends. In a robustness check, the authors use a regression discontinuity design to compare districts just above and just below the literacy threshold for receiving the program, yielding results that are significant and similar in size to the ITT estimates. Hong and Sarr (2012), Hahn et al. (2015), and Heath and Mobarak (2015) all evaluate the impact of the same female secondary school stipend (FSSP) in Bangladesh using a difference-in-difference strategy. Hahn et al. (2015) and Hong and Sarr (2012) both exploit the rule that only girls in rural areas were eligible for the stipend and compare education and marriage outcomes for rural vs. urban girls before and after the introduction of the FSSP. The main difference between the two is that Hong and Sarr (2012) also examine the earlier introduction of free secondary education for girls. Both papers find large effects on education and age of marriage from the FSSP (an increase in age of marriage of 0.6 to 2.3 years). It is worth noting two caveats to these results. First, Hahn et al. (2015) find the program has its largest effect on primary school enrollment which is odd for a secondary school program. More importantly, the raw age of marriage data suggests that the changing gap between rural and urban ages of marriage is driven mainly by a collapse in the age of marriage of urban girls rather than an increase in age of marriage of rural girls. Heath and Mobarak (2015) examine the impact of the FSSP in periurban areas (still classed as rural) by comparing girls who reached secondary school just before the FSSP and just after the FSSP and find no statistical difference. They note that the gender gap on education was closing long before the introduction of the FSSP and that there was no trend break in the steadily increasing secondary school enrollment rates in the country as a whole. A key difference with the other FSSP papers is that they are not exploiting geographic differences in eligibility (which could confound program impact with differential trends in rural vs. urban areas), and compare girls in a much shorter age window around the introduction of the program.

In summary, we conclude that there is limited reliable evidence on the impact on child marriage of cash or noncash incentives that are conditional on education especially in the South Asian context.

Evidence on the Impact of Incentives Conditional on Avoiding Child Marriage

The evidence on the impact on child marriage of cash or noncash incentives that are conditional on marriage is even more limited. Our search revealed four studies on such programs, (Krishnan et al., 2014; Nanda et al., 2014; Sinha and Yoong, 2009; Erulkar and Muthengi, 2009) three of which evaluated one of the longest running programs, Apni Beti Apni Dhan (ABAD) in India, (Krishnan et al., 2014; Nanda et al., 2014; Sinha and Yoong, 2009) and one evaluated a two-year intervention in Ethiopia (Erulkar and Muthengi, 2009). Only the study on the intervention in Ethiopia was published in a peer reviewed journal. Erulkar and Muthengi (2009) evaluate Berhane Hewane, a program that provided girls with a combination of group formation and study support and offered parents a financial incentive conditional on the girl still being unmarried at the end of the intervention. The authors find a significant effect on marriage; however, with only one control and one treatment community, the quality of the paper is too low for inclusion in our sample. The State of Haryana introduced the

ABAD program in 1994 with payments to mothers of girls at the time of their birth. The girl is also given a bond with a guaranteed payout of Rs 25,000 redeemable on her 18th birthday as long as she remains unmarried until that point. Thus the first payments tied to marriage did not take place until 2012. Additional bonuses are linked to the level of education the girl has achieved. Eligibility is limited to the first, second, and third children in a family that is below the poverty line and of specific castes. Because payments tied to marriage were only paid out for the first time in 2012, most studies have only examined the impact of the program on education, not marriage. An exception is Krishnan et al. (2014). However, the methodology used raises serious selection bias concerns. The authors compare ages of marriage for girls born in Haryana with daughters-in-law who have married into Haryana from other states, which did not have the ABAD program. They find that girls from Haryana have a higher age of marriage than daughters-in-law in Haryana who originally came from other states and attribute this difference to the program. However, girls marrying locally and those marrying into another state are likely to be very different on many dimensions and it is impossible to disentangle the effect of these factors on age of marriage from the impact of the program. Nanda et al. (2014) in contrast compared education outcomes for those eligible and ineligible for the ABAD program. However, they do not state how they determine eligibility. If, for example, eligibility is determined by whether mothers registered their daughter for the ABAD scheme at birth (as is required to claim the marriage contingent payment) then there would be substantial selection bias which would invalidate the comparison.

The highest-quality study of ABAD is Sinha and Yoong (2009). The authors utilize details of the eligibility criteria of the program to predict who is eligible. They use detailed asset information from surveys to predict whether a family is below the poverty line along with information on birth order and caste to determine if a girl is eligible for the program. The authors find no evidence that ABAD increased the likelihood of girls attending school although conditional on first attending, they were more likely to continue. Again there is a concern about differential trends as those eligible girls have lower initial schooling outcomes and thus might have experienced faster growth than ineligible girls.

Empowerment and Livelihood Programs and Age of Marriage

There were two high-quality studies of the effect of empowerment or livelihood programs on age of marriage or cohabitation or teenage pregnancy (Bandiera et al., 2014; Buehren et al., 2015). One was a working paper of a large clustered randomized control trial of BRAC's ELA centers in Uganda (Bandiera et al., 2014). Girls in communities with ELA centers were more likely to practice safe sex and were less likely to have had a pregnancy than those in control areas. The BRAC program was developed at the same time and along similar lines to the Kishoree Kontha program which we evaluate. One key difference between the studies is context: Bandiera et al. (2014) studies a program in Uganda where marriage and sexual onset are less determined by parents and are more in the control of girls than is the case in Bangladesh. Other differences are that: BRACs ELA centers are permanent while the KK program only lasts for six months; BRAC centers are run by adults while KK was run by peer educators; and BRAC centers provide vocational training while KK provided more general life skills training. Another empowerment program in Tanzania and modeled after the BRAC empowerment program in Uganda did not have significant effects on education or marriage outcomes (Buehren et al., 2015).

Below is a list of all studies included in the literature review above. Presented are authors, interventions evaluated, country of intervention, publication type, methodology and outcomes. The last two comments list the quality of the study (low, medium, high) as well as the order of reference in the bibliography to the literature review.

Table 2.1: Studies considered in literature review

Authors	Program Evaluated	Country	Type of Publication	Methodology	Outcome	Quality	Reference
Krishnan et al.	Apni Beti Apna Dhan	India	Peer reviewed	Local vs out-of-state girls	Education, marriage	Low	6
Baird et al.	CCT and UCT	Malawi	Peer reviewed	RCT	Schooling, pregnancy, marriage	High	7
Handa et al.	CT-OVC	Kenya	Peer reviewed	RCT	Pregnancy, marriage	Low	8
Duflo et al.	Free School Uniforms	Kenya	Peer reviewed	RCT	Schooling, pregnancy	High	9
Alam et al.	FSSP	Pakistan	Academic WP	Program vs. non-program states	Education, marriage	Low	10
Hong and Sarr	FSSAP	Bangladesh	Academic WP	Difference-in-difference	Education, marriage	Medium	11
Hahn et al.	FSSAP	Bangladesh	Academic WP	Difference-in-difference	Education, marriage	Medium	12
Angrist et al.	Vouchers to private schools	Colombia	Peer reviewed	RCT	Education, cohabitation	High	13
Heath and Mobarak	FSSAP	Bangladesh	Peer reviewed	Difference-in-Difference	Education, marriage	Medium	14
Nanda et al.	Apni Beti Aphna Dhan	India	Short report	Eligible vs ineligible	Education	Low	15
Sinha and Yoong	Apni Beti Apna Dhan	India	Academic WP	Instrument for eligibility	Education	Medium	16
Erulkar and Muthengi	Berhane Hewan	Ethiopia	Peer reviewed	Matched communities	Education, marriage	Low	17
Bandiera et al.	BRAC ELA centers	Uganda	Academic WP	RCT	Pregnancy, safe sex practices	High	18
Buehren et al.	BRAC ELA centers	Tanzania	Academic WP	RCT	Education, marriage	High	19

3 Survey Questionnaire and Variable Construction

Extract from endline questionnaire

SL	Question and Hints	Codes	Relevance
1.	Enter Enumerator ID		
2.	Enter Team ID		
3.	Enter Household ID		
4.	Re-enter Household ID		
5.	Which girl are you asking questions about?	1=Name: id: (Options prefilled) 2= Name: id: 3= . . .	
6.	Is the girl name listed in the pre-fill information correct?	1=Yes, the name is correct. 2=No, the name is another family member. 3=No, the name is entirely incorrect.	
7.	Please enter the correct girl name		If Q.6 is 2 or 3
8.	Is the father name listed in the pre-fill information correct?	1=Yes, the name is correct 2=No, the name is another family member 3=No, the name is entirely incorrect	
9.	Please enter the correct father name		If Q.8 is 2 or 3.
10.	Is the Para location listed on the pre-fill information correct? Hint: If the pre-fill data is missing, select “No”	1=Yes 2=No	
11.	Please enter the correct Para location		If Q.10 is 2.
12.	Is the Bari name listed in the pre-fill information correct? Hint: If the pre-fill data is missing, select “No”	1=Yes 2=No 96=Don’t know	
13.	Please enter the correct Bari name		If Q.12 is 2.
Note	Enumerator please note: Talk to the girl’s mother. If the girl’s mother is not available, talk to her father. If her father is also not available, ask “which household member, who is currently at home, would know best about this girl?”, and then continue with the consent form and survey.		

CONSENT (If Q.2 is 1 or 4)

Assalamu Alaikum/Adab,

My name is ______________. I have come from Mitra & Associates. Our organization is conducting a survey on behalf of the Abdul Latif Jameel Poverty Action Lab (J-PAL) at the Massachusetts Institute of Technology and Innovations for Poverty Action (IPA), two U.S.-based non-profit organizations dedicated to finding innovative solutions to development issues in various countries. The purpose of the study is to learn about the education, health, marriage, and economic activity of young women and adolescent girls. A few years ago we interviewed your household and the adolescent girls in it. If you agree, we would like to ask similar questions again now. There is no risk for you if you decide to participate in the study. There are no benefits either. I am now going to tell you a little bit more about the survey and ask if you and your household would be willing to participate.

Participating in this survey is totally up to your wish and you have all the freedom to not participate in it. If you want, you can stop the interview at any time, without penalty or consequences of any kind. If you volunteer to participate in this study, we would ask you for the following:

Information about the educational status, economic activity, as well as personal & health related matters of the girls and women aged 16-24 that we interviewed last time.

Information about the educational status and economic activity of the husbands of any girls who have married.

Information that would allow us to get in touch with any of the girls or young women who are no longer living with you.

The interview with you will last about 25 minutes. It will include some sensitive questions about the girl's marriage and children.

Your family or the people of your area will not be inconvenienced in any way from your participation or from allowing your children to participate in the survey.

The information provided by you will remain strictly confidential. Your name will never be published in any report. The published result of the survey will be the compilation of information of thousands of people and the information given by any specific person cannot be identified from there. In order to ensure that this survey is being administered appropriately, a short audio recording may be collected during your interview. The recording will be securely reviewed internally and deleted within 30 days of collection. No one outside of our quality control team will access the recording. You can ask any questions that you have about the study now. Do you have any questions? You can agree to your and your family's participation in the study either by informing me now. If you feel that anyone working on this survey is doing anything you are not happy about or if you have any questions about the research as a respondent, then please contact XXX or XXX, Evaluation Coordinators, IPA at XXX and XXX. You can also call the Chairman of the Committee on the Use of Humans as Experimental

Subjects at M.I.T. at +1- 617-253 6787 in the United States.

Do you understand the procedures described above?

Have your questions been answered to your satisfaction?

Do you agree to participate in this study?

1=Yes.

2=No.

All following questions relevant when consent = 1

Section I – For all girls

SL	Question	Codes and Hints	Relevance
1.1.	Is this girl/woman living in the household?	1=Yes 2=No 91=This girl does not exist 99=The girl is no longer living	
1.2.	If no, where is she now?	1=At her husband's house (living separately from in-laws) 2=Migrated elsewhere for work 3=Migrated elsewhere for studies 4=At her in-laws' house 96=Don't know 98=Other	If Q.1.1 is 2.
1.3.	How old is she?	In years (Enter -999 if response is "Don't know")	If Q.1.1 is 1 or 2.
1.4.	Is she still in school/college?	1=Yes 2=No 96=Don't know	If Q.1.1 is 1 or 2.
1.5.	What class is she in?	Select option from Education code	If Q.1.4 is 1.
1.6.	What is the highest class she passed?	Select option from Education code	If Q.1.1 is 1 or 2.
1.7.	What is her current marital status?	1=Married 2=Single, never married 3=Widowed 4=Divorced 5=Separated 6=Abandoned 96=Don't know 98=Other	If Q.1.1 is 1 or 2.
1.8.	If divorced / widowed / separated / abandoned / other – how long after her marriage did this happen?	Enter number of years/months/weeks/days 96=Don't know	If Q.1.7 is 3 or 4 or 5 or 6.

Section II – For girls who have married.

Section relevant when Q. 1.7 is 1 or 3 or 4 or 5 or 6

Note	Enumerator please note: For the questions to follow: If the girl is currently single, but has married before, ask about her most recent marriage. If the girl has married more than once, and is currently married, ask about her current marriage. Note: The questions in this section are not related to her fiancee. Do not enter information on her fiancee.

SL	Question	Codes and Hints	Relevance
2.1.	How long ago did she marry?	Enter number of Years/Months/Weeks/Days 96=Don't know	
2.2.	Is there a marriage certificate in the household?	1=Yes 2=No 96=Don't know	
2.3.	If yes, may I see the marriage certificate so that I can note the exact date?	1=Yes 2=No	If Q.2.2 is 1.
2.4.	Enumerator: Note date of marriage from certificate	Enter DD/MM/YYYY in the answer section	If Q.2.3 is 1.
2.5.	Is this her first marriage?	1=Yes 2=No	
2.6.	If no, how many times has she been married before? (Enumerator: Exclude current marriage)	Enter number of marriages	If Q.2.5 is 2.
2.6.a.	How long ago did she first marry?	Enter number of Years/Months/Weeks/Days 96=Don't know	If Q.2.5 is 2.

2.7.	Who arranged her marriage? Who was the medium involved? (Select multiple)	1=Professional *Ghatak/Raibaar* 2=Girl's Parents 3=Girl herself / love marriage 4=Neighbour 5=Relative 6=Teacher 7=Boy's Parents 96=Don't know 98=Other	
2.8.	How many marriage proposals did she get before you found this match? (Enumerator: Omit the current marriage. If girl has married more than once, ask for number of proposals BEFORE HER FIRST MARRIAGE ONLY.)	Enter number of proposals -999=Don't know	
2.9.	Did she go to her husband's house immediately after she got married?	1=Yes 2=No 96=Don't know 91=Not yet left for husband's place	
2.10.	How long after the marriage did she go to her husband's house?	Enter number of years/months/weeks/days 96=Don't know	If Q.2.9 is 2.
2.11.	Has she ever been pregnant?	1=Yes 2=No 96=Don't know	
2.12.	Is she pregnant at present?	1=Yes 2=No 96=Don't know	If Q.2.11 is 1.
2.13.	Has she ever given birth?	1=Yes 2=No 96=Don't know	If Q.2.11 is 1.

2.14.	If yes, how many live births has she given, in total?	Enter number of live births -999=Don't know	If Q.2.13 is 1.
2.15.	Why has she not given birth? (Select multiple)	1=Miscarriage 2=Abortion 3=Still born 96=Don't know 98=Other	If Q.2.12 is 2 and Q.2.13 is 2.
	Questions 2.16 to 2.18 relevant when Q. 2.13 is 1 and Q. 2.14 is not -999; Repeat for each child		
2.16.	How old is the child?	Enter number of years/months/weeks/days 96=Don't know 99=Child is not living	
2.17.	If the child is not living, how old was the child when he/she died?	Enter number of years/months/weeks/days 96=Don't know	If Q.2.16 is 99.
2.18.	Where did she give birth to the child?	1=Govt. Hospital 2=Govt. Health Center/Community Clinic 3=Govt. Dispensary 4=Private Hospital 5=Mission Hospital/NGO Hospital/Clinic 6=Private Clinic 7=Own house/Relative house/neighbour house 96=Don't know 99=Other	

Confidence Question

SL	Question and Hints	Codes
1.	Enumerator: Do you think the respondent was able to answer questions accurately? To the best of your judgment, how confident are you of the responses you got on this survey?	1=Very confident – Could answer almost all questions without hesitation 2=Somewhat confident – Hesitated a little, but was able to answer most questions 3=Unclear – Frequently hesitated and struggled, but answered some questions promptly 4=Definitely unsure – Frequently said "Don't know", struggled most of the time to come up with answers

Interview Result

SL	Question and Hints	Codes
1.	Interview Result:	1=Interview completed 2=Partially completed 3=Refused to answer

Education Code

0=Has not passed any class
1=Class One
2=Class Two
3=Class Three
4=Class Four
5=Class Five
6=Class Six
7=Class Seven
8=Class Eight
9=Class Nine
10=Class 10
11=SSC / Equivalent / Dakhil
12=HSC/First year
13= HSC/2nd year
14= HSC/ Equivalent /Alim
15=Vocational training
16= Honours 1st year
17= Honours 2nd year/pass
18= Honours 3rd year/Bachelor pass/Fazil
19=Bachelor Honours/BSc/Masters Preliminary
20=M.A/M.Sc/M.com/MSS/MBA
21=M.Phil
22=PhD/Post MBBS
50=Hafizi/Religious education
96=Don't know
98=Other

Table S2: Construction of variables

Type	Variable	Survey Question	Values/Construction
Marriage Outcomes:	Marriage age	Q. 2.1: How long ago did she get married?, Q.2.6 for girls who were married before.	Predicted age using baseline age - response (years+months+weeks+days), Minimum age=9
	Child marriage		1 – married, marriage age <18 0 – unmarried, marriage age>=18
	Under 20		1 – married, marriage age <20 0 – unmarried, marriage age>=20
	Under 16		1 – married, marriage age <16 0 – unmarried, marriage age>=16
Childbearing Outcomes:	Teenage pregnancy	Q. 2.16: How old is the child? Q. 2.18: If the child is not living, how old was the child when he/she died?	Age at 1st birth: Predicted age using baseline age – maximum response, Minimum age=11; Teenage pregnancy: 1 – given birth, age at 1st birth <20 0 – given birth, age at 1st birth >=20
Education Outcomes:	Still in school	Q 1.4: Is she still in school/college?	1 – Still in school. 0 – Not in school.
	Last class passed	Q1.5: What class is she in? Q1.6: What is the highest class she passed?	Highest class passed for girls out of school, current class – 1 for girls in school.
Baseline Covariates:	Mother education	Highest class passed	Response by female HH head or spouse of HH head
	HH size		Number of members in baseline household.
	Girl-to-boy ratio		Ratio of girls age 14 to 16 to boys age 16 to 18 in community.
	Older unmarried sister		1 – If there was an older unmarried sister living in the HH at baseline. 0 – No older unmarried sister living in the HH.

4 Merging Accuracy

Table 4.1: Merging accuracy, percentage of observations that have Levenshtein distances of more than 2, 3, or 4 with their merged baseline girls, which means that more than 2, 3, or 4 single-character edits would be required to move from the baseline name to the endline name after phonetic cleaning. Girls age 15-17 and unmarried at distribution start

	Empowerment (%)			**Incentive** (%)			**Empow.+Incen.** (%)			**Control** (%)		**Total** (%)	
Parents' Survey													
N		5,119			2,349			2,659			5,333		15,460
	Mean	S.D.	Diff.	Mean	S.D.	Diff.	Mean	S.D.	Diff.	Mean	S.D.	Mean	S.D.
Levenshtein distance> 2	1.5	12.2	-0.4	2.6	15.9	0.7	1.7	13.0	-0.2	1.9	13.7	1.8	13.5
Levenshtein distance> 3	1.2	10.8	-0.4	2.1	14.4	0.6	1.4	11.6	-0.2	1.6	12.4	1.5	12.1
Levenshtein distance> 4	0.8	8.9	-0.2	1.4	11.6	0.3	0.9	9.5	-0.1	1.0	10.1	1.0	9.9
Subsample													
N		790			452			494			841		2,577
Levenshtein distance> 2	13.0	33.6	0.1	15.7	36.4	2.8	9.3	29.1	-3.6*	12.9	33.5	12.7	33.3
Levenshtein distance> 3	9.9	29.9	0.3	12.6	33.3	3.0	6.3	24.3	-3.3*	9.6	29.5	9.6	29.0
Levenshtein distance> 4	6.4	24.5	0.6	7.1	25.8	1.3	4.4	20.5	-1.5	5.9	23.5	6.0	23.3

Data are means, standard deviations, and differences from comparisons with Control group. Differences are rounding to one decimal place. Significance levels are * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.

We analyze all 15,464 girls in our sample. Girls are merged to baseline observations based on household ID, name, birth order, age, and household characteristics. We list error rates based on names in Table 4.1.

Error rates based on household and names are very low, suggesting a high success rate in tracking the baseline girls. Name error rates are marginally unbalanced across treatment arms when looking at the percentage of names that would require more than two Levenshtein movements, that is the minimum number of single-character edits to move from the baseline name to endline name after phonetic cleaning, with slightly higher error rates among girls eligible for the incentive. However, these differences in merging success are not related to treatment. We count the distance as 0 if one name completely contains the other. In addition, girls with non-matching names may still be appropriate matches as it is not unusual for girls to call themselves different names in different settings in Bangladesh (e.g. with family, friends, and outsiders).

5 Randomization

Prior to randomization, the list of communities was organized in two steps:

1. Communities were organized by the number of girls age 10 to 19 at baseline.

2. Size tiers of communities were determined and communities then ordered by unionIDs and size tiers, whereby the order of communities within union and size tiers was random.

First, the number of multiples of 6 was determined per union as this was the number of randomizations to be performed per union (each treatment status was related to one number with the empowerment and control arms being assigned two numbers).

Then, in each union, the treatment status of the first community was randomized. The treatment status of the following communities was assigned in the sequence of 1 to 6 (e.g. if the first community was randomly assigned treatment 3, the subsequent communities were assigned treatments 4,5,6,1 etc.).

Lastly, all remainder communities in excess of the multiples of 6 were ordered by unions and tiers and the treatment status of the first community assigned by randomization and of all subsequent remainder communities by filling the sequence 1-6.

6 Take-up Adjusted for Crossovers

Table 6.1: Take-Up, calculated from monitoring data, adjusted for crossovers. KK: Girls age 10-19 in empowerment villages, Incentive: Unmarried girls age 15-17 in incentive villages at distribution start

Treatment Group	**KK Enrollment** (%)		**KK Attendance, Unconditional** (%)		**Oil Take-up** (%)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Empowerment	73.3	8.9	55.8	8.6	.	.
Incentive	.	.	.	.	62.6	16.0
Empowerment+Incentive	78.8	6.7	61.4	7.6	69.4	14.3
Any Empowerment	75.2	8.6	57.7	8.7	.	.
Any Incentive	.	.	.	.	66.0	15.5

Table 6.2: Take-Up, self-reported. Any empowerment includes girls in empowerment and empowerment plus incentive treatment groups. Any incentive includes girls in the incentive and empowerment plus incentive treatment groups. Girls age 15-17 and unmarried at distribution start

Treatment Group	**Attended at least 1 KK session** (%)		**Member of KK** (%)		**Oil Take-up** (%)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Empowerment	47.0	49.9	40.5	49.9	0.6	49.9
Incentive	29.7	45.7	23.3	45.7	62.8	45.7
Empowerment+Incentive	71.3	45.3	64.8	45.3	69.4	45.3
Control	12.3	32.8	8.8	32.8	1.0	32.8
Any Empowerment	56.1	56.1	49.6	50.0	.	.
Any Incentive	.	.	.	.	66.2	47.3

7 Cost-Benefit and Cost-Effectiveness Analyses

We conducted cost-benefit and cost-effectiveness analyses comparing various approaches to reducing the incidence of early marriage. Drawing on our literature review (supplementary online appendix 2), we include only interventions for which there are medium to high-quality evaluations which tracked child marriage or marriage age as an outcome. Including the conditional incentives to delay marriage in Bangladesh, we consider 6 programs from South Asia, Sub-Saharan African, and Bangladesh which demonstrate impacts on child marriage rates or girls' age of marriage.

7.1 Interventions Included in Analyses

<u>Intervention 1: The Female Secondary School Stipend Assistance Program in Bangladesh</u>

Bangladesh introduced the Female Secondary School Stipend Assistance Program (FSSAP), a large-scale education promotion program, in 1994 to make secondary education free for girls in rural areas. The program aimed to address the gender gap in secondary education by encouraging more girls to complete secondary education. In addition to free schooling, the FSSAP paid a small stipend to eligible girls conditional on their enrollment, a minimum 75% attendance rate in school, a minimum 45% average on annual exams, and remaining unmarried. The level of the stipend varied by grade and by year. The program covered more than two million girls each year and was the Bangladesh government's most prominent education program through the 1990s and 2000s (Hahn et al., 2015). The program later was changed to include stipends for poor boys in addition to girls. We only consider the first iteration of the FSSAP in our analysis. As addressed in the literature review (supplementary online appendix 2), several papers evaluate the impact of the FSSAP in Bangladesh using a difference-in-difference strategy and find impacts on age of marriage ranging from zero to an increase of 2-3 years (Hong and Sarr, 2012; Hahn et al., 2015; Heath and Mobarak, 2015).

<u>Intervention 2: Vouchers for Private Education in Colombia</u>

In 1991, the Colombian government established a voucher program for low income students to attend private schools as a way to rapidly expand secondary school access despite limited public secondary schools. The Programa de Ampliación de Cobertura de la Educación Secundaria (PACES) was one of the largest voucher programs to date, providing over 125,000 students from poor urban neighborhoods with vouchers that cover more than half of the cost of private secondary schools in Colombia. PACES vouchers covered the average tuition of low-to-middle cost private schools in Colombia's largest cities. The vouchers were available for both boys and girls and distributed by random lottery within the pool of eligible applicants. The random allocation of the vouchers has allowed for a series of evaluations to establish rigorous evidence of the impacts of the program (Angrist et al., 2006, 2002; Bettinger et al., 2010, 2016). The vouchers increased test scores for girls and led to a decrease in cohabitation.

Intervention 3: Free School Uniforms in Kenya

While the Kenyan government eliminated school fees in 2003, students are generally still required to purchase and wear uniforms to attend school. The Child Sponsorship Program, a project of ICS-Africa, implemented a non-cash incentive intervention in Busia, Kenya which provided free school uniforms to primary school students. The uniforms were meant to decrease financial barriers to schooling and raise attendance for both boys and girls. Schools were randomly assigned to participate in the program which allowed for causal inferences of the impacts of the intervention. In total, 83 schools received the stand-alone school uniform subsidy with an average of 29.3 eligible girls per school (Duflo et al., 2015). The program increased schooling and decreased marriage rates.

Intervention 4: Empowering Adolescent Girls in Uganda

From 2008-2010, a large cluster randomized trial evaluated BRAC's Empowerment and Livelihood for Adolescents Program (ELA) in Uganda. The program worked with BRAC's permanent centers in communities to provide life skills and vocational training for girls aged 14-20 through adolescent development clubs and sessions led by young female mentors. The clubs were open five afternoons a week after school and covered issues of sexual and reproductive health, menstruation, pregnancy, STIs and HIV/AIDS awareness, family planning, and rape. Additionally, trainings provided information on conflict resolution and legal standards regarding bride-price, child marriage, and domestic violence. Vocational skills training focused on small business development, including courses on tailoring, computing, and dancing. The clubs served as a center of recreation for the girls and provided a safe space in which they could meet and privately discuss their problems. ELA centers led to decreases in underage marriage and cohabitation and significantly reduced childbearing. Additionally, girls in communities with ELA centers were more likely to practice safe sex and were less likely to have had a pregnancy (Bandiera et al., 2014).

Intervention 5: Unconditional Cash Incentives for Girls in Malawi

Between 2008 and 2009, researchers performed a randomized evaluation of an unconditional cash incentive (UCT) program for girls in Malawi. The study took place in southern Malawi's Zomba district in both the large urban center, Zomba city, and many surrounding rural and semi-rural communities. Girls and their families received monthly stipends over the course of two school years. The incentives ranged from \$4 to \$10 plus the cost of school fees. Girls receiving the cash stipends were 8 percentage points less likely to be married and 7 percentage points less likely to be pregnant than girls in the comparison group. A parallel study of a cash incentive conditional on 80% school attendance found no significant effects on marriage or pregnancy outcomes, but did find improvements in school enrollment, attendance, and test scores. The UCT program may have been more effective in delaying marriage and childbearing because it allowed girls who dropped out of school to support themselves without relying on a husband (Baird et al., 2011).

Intervention 6: Conditional Financial Incentives in Bangladesh

In collaboration with Save the Children (USA), a large clustered randomized trial in southern Bangladesh examined a conditional stipend program which encouraged parents not to marry their adolescent daughters before the legal age of consent. The program distributed cooking oil to girls aged 15 through 17 and confirmed to be unmarried. The oil was distributed every four months between April 2008 and August 2010 with monitors checking the marital status of the girls before each distribution. ELA centers led to decreases in underage marriage and cohabitation and significantly reduced childbearing. Additionally, girls in villages with ELA centers were more likely to practice safe sex and were less likely to have had a pregnancy.

7.2 Methodology

Benefits of delayed marriage are calculated based on the cumulative education wage premium for girls eligible for each program. We assume girls start working at age 17.64, the median age of marriage for girls in the control group of the conditional incentives in Bangladesh evaluation, and continue working until they are 60. We assume that wage returns to education are constant across their working life, and that the returns to years of secondary education are equal for women in and out of the workforce. We assume that extra education delays girls' entries into the workforce, and that they begin working immediately after finishing their studies. To consider all interventions in the same time frame, we consider all interventions had they started in 2008.

To estimate girls' income, we use the following equation:

$$income = \gamma \times (income\ for\ rural\ girls) + (1 - \gamma) \times (income\ for\ urban\ girls)$$

Where γ is equal to 0.7 to reflect the relative proportion of rural to urban girls from the conditional incentives evaluation in Bangladesh. All income streams are thus calculated twice, once for rural areas and once for urban areas, and then weighted.

To estimate the wage premium benefits of the program, we use the estimated education wage premium from Asadullah (2006) to determine the estimated income for girls in each year of each program. He uses a modified Mincer equation to estimate girls' expected wages in each year of their working life as a function of education and experience, which we modify to calculate a girl's wage in each year:

$$income_t = (1+\rho)^t \times e^{\alpha\ +\beta_1(years\ of\ schooling)+\beta_2(experience_t)+\beta_3\left(experience_t{}^2\right)+\beta_4(female)+\beta_5(rural_i)+\varepsilon}$$

Where α is the log of mean monthly income in 2008 BDT for an urban man in Bangladesh with no education or experience. We estimate α from Asadullah's 2000 estimate of hourly wages assuming 8 hour work days and 25 working days per month as well as by inflating to 2008 BDT by using the average BDT inflation rate between 2000 and 2014, 5.90%, (World Development Indicators, 2016a) and ρ, the mean GDP per capita growth rate in Bangladesh 2000-2014, 4.26% (World Development Indicators, 2016b). Experience is calculated as the number of years since a woman entered the workforce. t is the year of the analysis and we account for growth in wage levels over time. i indicates whether a girl resides in a rural area. All interventions are considered as beginning in 2008, the first year of the oil incentive program, which we use as the base year for our calculations. Girls enter the workforce in 2010, defined as the year they reach median age of marriage, adjusted for additional education induced by each intervention.

To calculate the benefits for each year of an intervention, we first calculate estimated wages for a girl receiving the intervention and the estimated wages for a girl not receiving the program.

For girls receiving the program, income in each year is estimated as:

$$(1+\rho)^t \times e^{\alpha\ +\beta_1(median\ schooling+\lambda)+\beta_2(t-median\ marriage\ age-\lambda)+\beta_3(t-median\ marriage\ age-\lambda)^2+\beta_4(female)+\beta_5(rural_i)}$$

Where λ is the point estimate for the education benefit of a program. For girls not receiving the program, income in each year is estimated as:

$$(1+\rho)^t \times e^{\alpha\ +\beta_1(median\ schooling)+\beta_2(t-median\ marriage\ age)+\beta_3(t-median\ marriage\ age)^2+\beta_4(female)+\beta_5(rural_i)}$$

For the girls receiving the program, this takes into account the educational income premium from the additional years of schooling induced by the program as well as the loss of work experience from staying in school. As mentioned above, experience begins to accrue for all girls not in school after the median age at first marriage.

The present value of both costs and benefits of each program are then defined as follows:

$$Present\ Value\ = \sum_{t=0}^{\mathrm{T}} \frac{(annual\ program\ costs/benefits_t)}{(1+\theta)^t}$$

Where T is the number of years between the beginning of the intervention and the end of a woman's working life and θ is the social discount rate. We report results using a social discount rate of 5% and perform sensitivity checks using discount rates of 3% and 10%.

For all calculations, we first take the present value of the program cost and benefit streams in 2008 BDT as described above. For all costs not in BDT we first convert from local currency to BDT in the year of the intervention and inflate to 2008 BDT. We then inflate the streams to 2014 BDT using the average BDT inflation rate between 2000 and 2014 (World Development Indicators, 2016a). Finally, we convert the streams from 2014 BDT to 2014 USD using an exchange rate of 77.64 BDT per USD (World Development Indicators, 2016c).

Benefits Calculations

We define the benefits each year as the estimated income for a girl receiving the intervention minus the estimated income a girl in the program would have received without the education benefit of the program. This is the annual education benefit we expect girls to receive from having been induced to study longer by each intervention. As detailed above, the estimated income for a girl receiving a program is:

$$(1+\rho)^t \times e^{\alpha\ +\beta_1(median\ schooling+\lambda)+\beta_2(t-median\ marriage\ age-\lambda)+\beta_3(t-median\ marriage\ age-\lambda)^2+\beta_4(female)+\beta_5(rural_i)}$$

From this we subtract the counterfactual wages for a girl who participated in the program, but did not receive the education benefit:

$$(1+\rho)^t \times e^{\alpha\ +\beta_1(median\ schooling)+\beta_2(t-median\ marriage\ age-\lambda)+\beta_3(t-median\ marriage\ age-\lambda)^2+\beta_4(\mathrm{female})+\beta_5(rural_i)}$$

This gives us the benefit per eligible girl in any given year. The annual benefit is defined as the above term multiplied by the total number of girls eligible.

For the cost-benefit analyses, we also consider the value of stipends or incentives as benefits, discounted to 2008 and inflated to 2014 USD as described above.

Note that wage premiums are based on wages of those in the labor force with monetized wages. The assumption behind Mincer equations and our estimates is that education increases productivity as much for women not earning a wage (including those working in the household) as it does for women working for a wage outside the household. Our results are likely to be sensitive to this assumption, but the assumption has the same impact on all programs equally. Cost-effectiveness calculations using only costs to implementers, however, do not require reliance on this assumption and simply express additional years of education in terms of spending on a given program.

Estimate of Education Impact from Delayed Marriage

For each program for which we have data on impacts on education, we calculate the educational benefits to delayed marriage as follows: First, we estimate the additional years of schooling per year of delayed marriage for the Conditional Incentives to Delay Marriage in Bangladesh. We then apply this conversion factor to the other interventions to estimate the implied additional years of schooling had the program taken place in southern Bangladesh. For studies in which we have age at first marriage as an outcome, this results in a conversion factor of 0.50 additional years of schooling for every additional year unmarried. Where we only have child marriage rates, we use a conversion factor of 0.02 additional years of schooling for every percentage point reduction in child marriage.

We take this approach for a number of reasons. First, the quality of the evaluations varies and the estimated educational returns to the program may not be equally reliable. Standardizing the assumed educational returns to delayed marriage from a recent rigorous evaluation helps normalize the quality of the estimates. Secondly, not all of the interventions took place in Bangladesh. We might expect the educational returns to delayed marriage to be quite different in Kenya and Colombia than in Bangladesh. Estimating the educational benefits using the educational returns from Bangladesh allows us to estimate the benefit-cost ratio of the other programs had they been implemented in Bangladesh.

Years of schooling and age of marriage are causally linked in both directions; delaying marriage leads to more schooling and more schooling leads girls to delay marriage. Many of the programs we consider in the CBA are principally intended to increase girls' educational attainment. In this CBA we are primarily interested in programs that reduce child marriage and seek to quantify the benefits of that reduction in child marriage through its impact on education. By applying a conversion factor derived from a program that primarily targeted age of marriage (and influenced age of marriage even for out of school girls) we may be disadvantaging the conditional incentive program at the expense of those programs more focused on education.

Cost Calculations

For all programs, we consider the costs to the beneficiary as well as the costs to the implementer. For the cost to the implementer, we consider actual program costs where available and estimate program costs elsewhere. Where monitoring costs for programs with conditional eligibility were not available, we consider the monitoring costs per girl per year of the oil incentive program multiplied by the number of eligible girls and years in the relevant intervention.

We also consider the opportunity cost of girls' foregone income over their entire working life from having fewer years of work experience. This cost is the difference in lifetime income for a girl due to having less experience from having been induced to stay in school. For each year, we calculate this cost by first estimating the income of a girl with median education and experience:

$$(1+\rho)^t \times e^{\alpha \ +\beta_1(\text{median schooling})+\beta_2(t-median\ marriage\ age)+\beta_3(t-median\ marriage\ age)^2+\beta_4(female)+\beta_5(rural_i)}$$

From this we subtract the counterfactual income of a girl with median schooling and fewer years of experience equal to the education effect of the relevant intervention:

$$(1+\rho)^t \times e^{\alpha\ +\beta_1(\text{median schooling})+\beta_2(t-median\ marriage\ age-\lambda)+\beta_3(t-median\ marriage\ age-\ \lambda)^2+\beta_4(female)+\beta_5(rural_i)}$$

This term is the same term subtracted from the wage of a girl in the treatment arm to calculate the benefits of the program and results in the foregone wages due to fewer years of workforce experience. The total costs of each program include the program specific costs to the implementer plus the income opportunity cost of education over a participant's working life. We calculate the net present value of the cost stream using the method described above.

Additional Assumptions for the Cost and Benefit Estimates for Specific Interventions

In addition to these general assumptions, intervention-specific assumptions are described in more detail below.

Intervention 1: The Female Secondary School Stipend Assistance Program in Bangladesh

Because we did not have access to administrative cost data, we estimate program costs as the costs of secondary school stipends plus the monitoring costs per girl from the oil incentive program. For the incentive cost of the stipends, we apply a incentive fee of Tk. 75 per wire as well as a cost of Tk. 1,000 to open a bank account. The stipend costs are estimated by using the number of total stipends per year, the amount of the average stipend per grade, and assuming equal distribution of stipends across grade years, and that the amount of the stipend remained constant for girls while they participated in the program. We use estimates of cohort size from several sources to estimate the number of girls eligible for the stipend each year (Hong and Sarr, 2012; Hahn et al., 2015; Raynor et al., 2006; Raynor, 2016). We assume equal distribution of girls per grade. This results in estimates of 8,728,387 eligible girls.

Intervention 2: Vouchers for Private Education in Colombia

To estimate the cost of the voucher program, we use average annual secondary education costs to the government per female scholarship winner (Angrist et al., 2002). The costs include the annual value of a PACES scholarship and the expenditure from scholarship costs for students who would have enrolled in private school adjusted for expenditures resulting from incentives from public to private schools and cost savings from reduced grade repetition.

Because researchers did not directly measure age at first marriage, we use the reported change in teen pregnancy as a proxy for changes in child marriage rates. Additionally, since the results come from a non-published intermediate paper, both the costs and benefit results may still change.

To establish the number of eligible girls we use the number of vouchers distributed, 90,000, divided by the take-up rate, 90%, for a full sample of 100,000 girls (Angrist et al., 2002). For each subsequent year, we use a take-up rate of 78%, the average re-enrollment rate for scholarship recipients (Angrist et al., 2002).

Since the scholarship program moved some students from public to private schools, there may have been a number of costs and benefits to the government which we do not include. For example, we do not count changes in tertiary education costs, loan subsidies, forgone tax revenue from VAT tax, changes in government revenue, nor forgone net government incentives through payroll taxes. Many of these costs are specific to the Colombian government context and these costs may not be relevant for replications of the program in Bangladesh. Additionally, we exclude those costs for better comparability to the other programs considered.

Intervention 3: Free School Uniforms in Kenya

For the Free School Uniforms in Kenya we use reported program costs including the cost of girls' school uniforms, NGO worker wages and NGO worker travel cost. These costs are all detailed on the JPAL website cost-effectiveness section.

Intervention 4: Empowering Adolescent Girls in Uganda

To estimate the costs and benefits of the BRAC Uganda program, we use extensive program cost data from Bandiera (2014). Reported costs for the 3,964 girls in the treatment group include, office space and equipment, program management and staff compensation, training and refresher course costs for adolescent leaders, club materials and rent, and the direct costs of financial literacy and livelihood trainings for girls. We exclude the country and branch office overheads reported in the paper to be consistent with cost calculations across programs.

Intervention 5: Unconditional Cash Incentives for Girls in Malawi

For the UCT in Malawi, we estimate the number of girls eligible for the program by multiplying the average number of girls per enumeration area (the unit of randomization) by the total number of enumeration areas. Costs include the fixed and variable costs to distribute the incentives, the cost of a parents' survey to establish the number of eligible girls, as well as the value of the incentives themselves.

Intervention 6: Conditional Incentives in Bangladesh

We have the most complete cost data for the conditional financial incentives to delay marriage in Bangladesh. Cost estimates include the costs of oil, monitors to confirm girls' marital status, transportation costs to deliver the cooking oil, and staff salaries of district point people, field officers, volunteers, and distribution workers.

The oil incentives were delivered through a food security program (Jibon-O-Jibika or JOJ) in the area that provided food incentives to pregnant and lactating mothers. JOJ's existing infrastructure led to cost-savings for the oil incentive program. We have estimated the program costs excluding the benefit from working with an existing distribution partner.

We only consider costs and benefits for girls who were 15 at oil distribution start because they received the full program, as would be the case if the program were scaled up. To estimate costs, we assume the costs for 15-year-old girls are proportional to the percentage of 15-year-old girls eligible for the program each year. We calculate all the costs, which is for girls aged 15 to 17, and then discount them to reflect the portion of girls who were 15 at the time of the oil distribution. This portion increases for each year as a result of girls aging out of the program. We consider costs and benefits for the 2,866 unmarried girls aged 15 in the 154 communities that were eligible for the oil incentive at the start of the oil distribution.

Net Present Value (NPV) per $1,000 Investment

We calculate the Net Present Value as the difference between the discounted benefits and the discounted costs of each program. Since the programs vary widely in terms of scale, the NPVs themselves are not directly comparable. The larger programs we consider have considerably larger NPVs, but only because they reached many more girls at scale. To aid in comparison of the programs, we divide the NPV by the total amount invested (costs to implementer and beneficiaries). We then present the NPV for each program in terms of NPV per $1,000 invested.

Benefit-Cost Ratio

We present a benefit-cost ratio for each of the programs. For these calculations we divide the total NPV of the benefits of a program by the NPV of its costs. All discounting and conversions are calculated as described above.

Cost-Effectiveness Analysis

In addition to the CBA, we provide estimates of the cost-effectiveness of each program in terms of a variety of outcomes including child marriages averted, additional years of schooling, and years of delayed marriage. These estimates are meant to give a sense of the relative efficiency of the programs at meeting particular outcomes without requiring the full set of assumptions of a CBA. To calculate the cost-effectiveness, we estimate the amount of a given outcome achieved by a given investment.

For example, for child marriages averted, we first multiply the point estimate of reduction in underage marriage by the number of girls eligible by the program. This gives us the total number of child marriages averted by the program. We then divide the present value of the costs of the program by the number of child marriages averted to determine the cost per child marriage averted. We then divide 1,000 by the cost per child marriage averted to express the figure in terms of returns to a $1,000 investment. For all interventions, we include both foregone income and implementation costs.

Limitations

Our estimates of the returns to education are based on a standard Mincer equation which compares earnings for those women in the workforce with different levels of education. Two key assumptions are necessary for this to reflect the gains to the economy of increases in education. First, it assumes the high wages of those women who are more educated are the result only of their education and not due to unobservables (such as motivation) which may be correlated with higher than average education. Second, it assumes that women who are not in the workforce but have had more education have an equal increase in productivity in the work they do at home as those who are in the workforce. If there is selection of more able or more motivated women into education or if education raises productivity less for those not in the labor force then our estimates will overestimate the Net Present Value of all the programs discussed here.

Causal evidence shows the effects of child marriage on educational attainment. However, as discussed above, corollary links form the bulk of the evidence on other effects of child marriage. We may imagine a range of long-term benefits for adolescent girls who delay marriage in terms of improved health outcomes, higher household decision-making power, or intergenerational well-being. The lack of causal evidence on these channels along with the difficulty of monetizing their benefits severely limits what can be included in a cost-benefit analysis. As a result, when monetizing benefits, we consider only the benefits of delayed marriage through the wage premium from increased years of schooling. We are therefore likely undervaluing the total benefits of each intervention.

Similarly, on the cost side, only the direct costs of each program and foregone wages are considered. We do not include costs associated with changes in the size or timing of dowries, for instance, or other less tangible benefits and costs that would require a large number of additional assumptions for which evidence is limited.

The availability of cost data varies across interventions. For the conditional incentives in Bangladesh, school uniforms in Kenya, and vouchers in Colombia, we have extensive information about the actual costs of delivering the interventions. However, for the FSSAP programs in Bangladesh, we have estimates solely on the cost of the stipends themselves and thus have to estimate other costs. For the BRAC Uganda intervention, we have extensive records of the program costs, but only a rough estimates of the number of girls eligible. This makes true comparisons of the efficiencies of the programs difficult and may lead us to overestimate the cost-effectiveness of the programs for which we have limited cost data. In particular, the high involvement of researchers in programs evaluated by RCTs may mean that cost data are collected more comprehensively making these programs appear costlier. In addition, those programs tested at small scale may have higher costs than they would if they were scaled up. We attempt to address this in at least one way by using the converted monitoring costs for programs for which we do not have monitoring cost estimates.

We are unable to take into account general equilibrium effects. Most importantly, if increases in education lead to a decline in the marginal return to education this would depress the benefits of all the interventions discussed here. Working in the other direction there may be social benefits to education not captured in Mincer regressions and even complementarities in the returns to additional education which would suggest our benefits are underestimates.

The dynamics of marriage are very different in South Asia than in Africa and Latin America where several of our comparative cases took place. Whereas parents in Bangladesh exercise significant control over their daughters' marriage-age decisions, in Malawi and Colombia girls have more influence on their own decisions. As a result, much of the evidence from regions outside of South Asia focus on teenage sexual activity or cohabitation rather than marriage. These differences make cross-comparisons difficult, and we may expect interventions to have very different results given the dynamics of each cultural context.

References

Andaleeb Alam, Javier Eduardo Baez, and Ximena V Del Carpio. Does cash for school influence young women's behavior in the longer term? evidence from pakistan. *World Bank Policy Research Working Paper Series*, 2011.

Joshua Angrist, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. Vouchers for private schooling in colombia: Evidence from a randomized natural experiment. *The American Economic Review*, 92(5):1535–1558, 2002.

Joshua Angrist, Eric Bettinger, and Michael Kremer. Long-term educational consequences of secondary school vouchers: Evidence from administrative records in colombia. *The American Economic Review*, 96(3):847–862, 2006.

Mohammad Niaz Asadullah. Returns to education in bangladesh. *Education Economics*, 14 (4):453–468, 2006.

Sarah Baird, Craig McIntosh, and Berk Özler. Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, qjr032, 2011.

Oriana Bandiera, Robin Burgess, Markus Goldstein, Niklas Buehren, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. Women's empowerment in action: Evidence from a randomized control trial in africa. *The London School of Economics and Political Science, Suntory and Toyota International Centres for Economics and Related Disciplines*, 2014.

Eric Bettinger, Michael Kremer, and Juan E Saavedra. Are educational vouchers only redistributive? *The Economic Journal*, 120(546):F204–F228, 2010.

Eric Bettinger, Michael Kremer, Maurice Kugler, Carlos Medina, Christian Posso, and Juan Esteban Saavedra. Educational, labor market, and fiscal impacts of scholarships for private secondary school: Evidence from colombia. 2016.

Niklas Buehren, Markus Goldstein, Selim Gulesci, Sulaiman Munshi, and Yam Venus. Evaluation of layering microfinance on an adolescent development program for girls in tanzania. *Working Paper*, 2015.

Esther Duflo, Pascaline Dupas, and Michael Kremer. Education, hiv, and early fertility: Experimental evidence from kenya. *The American Economic Review*, 105(9):2757–2797, 2015.

Annabel S Erulkar and Eunice Muthengi. Evaluation of berhane hewan: A program to delay child marriage in rural ethiopia. *International Perspectives on Sexual and Reproductive Health*, pages 6–14, 2009.

Sreela Das Gupta, Sushmita Mukherjee, and Sampurna Singh. *Knot ready: Lessons from India on delaying marriage for girls.* International Center for Research on Women (ICRW), 2008.

Youjin Hahn, Asadul Islam, Kanti Nuzhat, Russell Smyth, Hee-Seung Yang, et al. Education, marriage and fertility: Long-term evidence from a female stipend program in bangladesh. *Melbourne: Monash University*, 2015.

Sudhanshu Handa, Amber Peterman, Carolyn Huang, Carolyn Halpern, Audrey Pettifor, and Harsha Thirumurthy. Impact of the kenya cash transfer for orphans and vulnerable children on early pregnancy and marriage of adolescent girls. *Social Science & Medicine*, 141:36–45, 2015.

Rachel Heath and A Mushfiq Mobarak. Manufacturing growth and the lives of bangladeshi women. *Journal of Development Economics*, 115:1–15, 2015.

Seo Yeon Hong and Leopold Remi Sarr. Long-term impacts of the free tuition and female stipend programs on education attainment, age of marriage, and married women's labor market participation in bangladesh., 2012.

Anand Krishnan, Ritvik Amarchand, Peter Byass, Chandrakant Pandav, and Nawi Ng. "no one says 'no'to money"–a mixed methods approach for evaluating conditional cash transfer schemes to improve girl children's status in haryana, india. *International Journal for Equity in Health*, 13(1):1, 2014.

Susan Lee-Rife, Anju Malhotra, Ann Warner, and Allison McGonagle Glinski. What works to prevent child marriage: A review of the evidence. *Studies in Family Planning*, pages 287–303, 2012.

Anju Malhotra, Ann Warner, Allison McGonagle, Susan Lee-Rife, Cynthia Powell, Eva V Cantrell, and Reshma Trasi. Solutions to end child marriage. 2011.

B Mensch and CB Lloyd. *Growing up Global: The Changing Transitions to Adulthood in Developing Countries.* 2005.

Priya Nanda, Nitin Datta, and Priya Das. Impact of conditional cash transfers on girls' education. summary of research findings. 2014.

Jennifer Parsons and Jennifer McCleary-Sills. Preventing child marriage: Lessons from world bank group gender impact evaluations. 2005.

Janet Raynor. Email sent to: author, 2016.

Janet Raynor, Kate Wesson, and Milton Keynes. The girls' stipend program in bangladesh. *Journal of Education for International Development*, 2(2):1–12, 2006.

Nistha Sinha and Joanne Yoong. Long-term financial incentives and investment in daughters: Evidence from conditional cash transfers in north india. *World Bank Policy Research Working Paper Series, Vol*, 2009.

The World Bank World Development Indicators. Inflation, gdp deflator (annual %). http://data.worldbank.org/indicator/NY.GDP.DEFL.KD.ZG, 2016a.

The World Bank World Development Indicators. Gdp per capita growth (annual %). http://databank.worldbank.org/data/reports.aspx?source=2&series=NY.GDP.PCAP.KD.ZG, 2016b.

The World Bank World Development Indicators. Official exchange rate (lcu per us$, period average). http://data.worldbank.org/indicator/PA.NUS.FCRF, 2016c.