A Reassessment of 'Do Teenagers Respond to HIV Risk Information?'

Ana Catalano, Yanfang Su, Erru Yang

April 28, 2011

Introduction

In Pascaline Dupas' article, 'Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya' (*American Economic Journal*, January 2011), information on the relative risk of HIV infection (RR) by partners age is found to lead to a 28% decrease in teen pregnancy, a proxy for the incidence of unprotected sex. According to the author's analysis of self-reported sexual behavior, this is because teenage girls who have received the risk ratio information substitute away from older (riskier) partners toward younger partners. However, the experiment is complicated by the existence of another treatment, the Kenyan government's HIV prevention teacher training program (TT), which interacts with RR in some schools. Furthermore, the study is randomized at the school level, while RR treatment affects only 8th grade classes, resulting in a treatment 'spillover' problem across study and control cohorts which could bias difference-in-difference estimates.

In this article, we provide a better estimation strategy to account for both of these problems. A simple difference model allows us to exploit the study's randomization at the school level, comparing only those classrooms that were treated (all of which were in the 8th grade) to other 8th grade classes that were not treated. We then provide a falsifcation test for the 'placebo effect' theory of treatment spillover by comparing only 7th grade classrooms or only 9th grade classrooms (the control cohort, depending on the model). Finally, we use subgroup analysis in all of our models in order to disaggregate the true effect of RR from that of RR and TT together and RR conditional on TT.

We find that the RR information program is not as robust as Dupas' original findings suggest, and the

the treatment may indeed be harmful for boys. According to our analysis, the effect of RR on childbearing is only significant for the subgroup of RR conditional on TT, where it leads to a 77% decrease in teen pregnancy (we explain this somewhat surprising result by also considering the effect of the TT program alone, which increases teen pregnancy). The RR program also increases the number of both boys and girls who report that they have ever had sex; however, we find no evidence for Dupas' key finding that it decreases the age gap between the girl and her partner. Turning to boys, who are at greater risk for HIV with any increase in sexual behavior because girls of nearly all ages have higher relative HIV prevalence rates, the RR program increases the number of boys who report not only having sex, but having sex with multiple partners. We see no corresponding increase in the number who report using a condom. In sum, we find no evidence that the RR information program reduces risky sex for girls (as reported by the age gap between partner). In fact, we find the opposite: RR is actually harmful because it increases the overall incidence of sex, and it may be especially harmful for boys given their relative risk ratios.

The article proceeeds as follows: first, we discuss the methods and assumptions of the original paper, making several critiques which also apply more broadly to survey and experimental designs. Next, we propose an alternative way of estimating the effect of RR which allows us to exploit Dupas' study design and compare different subgroups, as well as implement a falsification test for the spillover theory. After showing that these subgroups are balanced, we make the case that a single difference model is powerful enough to estimate the effect of RR at the classroom level. We first present our findings for the pure effect of RR (where RR=1 and TT=0, compared to classrooms that got neither treatment). In the second part of our findings section, we also present findings for other subgroups: RR&TT, RR conditional on TT, and TT alone. Finally, we conclude with a discussion about implications for the RR program and comments on other similar experimental designs.

Background on the HIV Risk Information Program (RR) and the Kenya Teacher Training Progam (TT)

In this section, we give a brief overview of the HIV risk information program and the state teacher training program and how they were implemented, both as essential background for understanding the rest of the paper, and to show why we felt we could improve upon the model by using subgroup analysis.

The HIV risk ratio information treatment was implemented in 2004 in the 8th grade classrooms of 71

schools selected randomly among 328 schools. These 328 schools had been part of a randomized evaluation designed to test the effectiveness of the national HIV prevention teacher training program (TT). They were randomly selected and stratified by location, test scores, and student sex ratio.

Thus, there are two treatments. The teacher training program (TT) occurred first, in 2003, and involved teachers in 164 schools learning how to discuss HIV/AIDS issues in class, how to set up a health club in their school, and how to encourage HIV avoidance through active learning activities such as role plays.

The relative risk information program (RR) occurred after the teacher training program, in 2004. The motivation behind the RR program was to provide an alternative to curriculum which is largely limited to 'risk avoidance' (e.g. abstinence only) and provide 'risk reduction' information to students instead, so that they are aware of relatively low risk options. The original article is especially concerned with the program's effect on teenage girls, who are at greater risk for HIV if in a relationship with an older partner: "compared to relationships with teenage boys, cross-generational relationships pose a higher risk of HIV infection for teenage girls" (pp 2).

The RR program involved a trained NGO project officer visiting 71 schools (stratified by whether they had gotten the 2003 TT treatment), showing a video about the risks of cross-generational sex and sugar daddies and sharing detailed prevalence rates of HIV disaggregated by gender and age group in the nearby city. In the RR program, project officers did not volunteer information about condoms but did answer students' questions about them. See Figure 1 for a flow chart of the treatments.

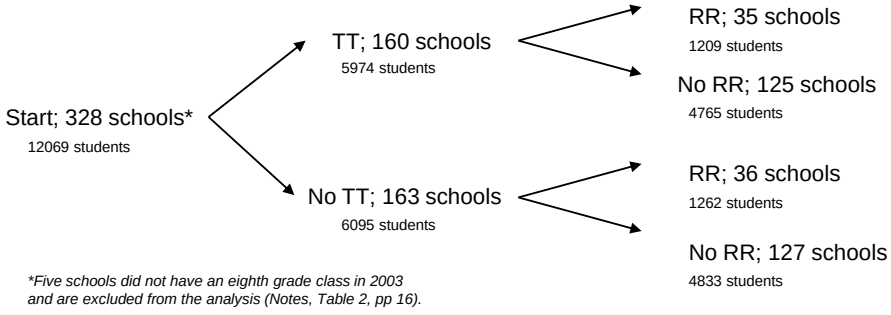

Figure 1: Flow chart of TT and RR treatments

The RR information program was given to 8th grade classrooms only, whereas the TT information program affected all upper primary classes in the school ("three teachers per primary school (typically, the math, English, and physical education teachers, rotating across upper primary classes) were trained", pp 8). Although is somewhat unclear which grades upper primary covers, we assume that at least grades 9 and 8

were affected.

Methods

Estimation Strategy for Multi-level Treatments in Experiments

It is the consensus that randomized experiments allow researchers to avoid many of the methodological complications that could otherwise come up in observational study to address problems such as omitted variable bias. However, when random assignment of treatments is clustered at multi levels (i.e., both school level and cohort level in Dupas's experiment), simple difference (SD) in means with clustering at the school level is biased. In this section we explore the methodological issues in experimental design with multi-level treatments, which is a key feature in Dupas' experiment, and introduce an alternative valid estimation strategy for data analysis. Furthermore, we employ the unique features of Dupas' study with multi-level treatments to conduct falsification tests, which falsify Dupas' theory about the potential negative spillover of RR treatment. Our paper contributes to the methodological discussion of multi-level treatment empirical framework.

Treatment and Control groups

Distinction of treatment group from control group is the first step we should carefully take in multi-level treatment design. For instance, in Dupas' experiment, the randomizations of RR and TT are both at the school level. However, Dupas only carried out the RR program for 8th graders in 2004 rather than for the whole school which is randomly chosen as the treated school. We believe that the reason that the RR treatment was not carried out for all the students at the selected school may have been due to financial or administrative constraints. One way around this problem is to randomly sample students within the school for treatment, allowing the researcher to conduct analysis of RR and TT treatments at school level. However, this study was designed instead to only treat 8th graders, whom are not randomly selected. There are two messages we should keep in mind:

(1) The 7th graders and 9th graders in the RR schools are untreated, just as students in non-RR schools are untreated, given no spillover and they should not be considered as "the treated"; (2) Simple difference

between the means of Y in the treated cohort and the untreated cohort is not the effect of RR because the treatment is not randomized. Both simple difference and difference-in-difference estimations in the original paper are thus biased due to neglecting this complexity in multi-level treatments.

First, the author applied simple difference in means of the treated schools and control schools to detect the effect of RR.

SD1=mean Y in the treated schools - mean Y in the control schools

Had all cohorts treated by the RR program, the effect of the RR could have been detected through SD_1. However, the observations in the treated schools include both the treated cohort and the untreated cohorts. Through SD1 regression, Dupas made an implicit assumption that all cohorts are treated by RR. Dupas mistakenly included the untreated cohorts as the treated in SD_1 regression, which effectively works as a downward bias, underestimating the treatment effect.

Second, the author also used a Difference-in-difference (DD) with interaction between a binary variable of treatment at school level and a binary variable of treatment at cohort level. This is actually a simple difference (SD) in means at the cohort level in the treated schools.

DD=SD_2=mean Y in the treated cohort in RR schools- mean Y in the control cohort in RR schools

However, as the treatment of RR is arbitrarily assigned to the 8th graders, the difference in means between the treated cohort and the control cohort is not necessarily due to the effect of treatment. Pre-existing characteristics between cohorts could bias-up or bias-down the estimate of RR effect. Besides, spillover effect is another concern to bias the estimate, which was expressed by Dupas as well.

Four-group Balance check and Simple Difference

We propose a simple and clear estimation strategy for two-treatment experiment.

All the schools enrolled in Dupas' study will be randomly selected into four groups: (1) both RR and TT; (2) only RR without TT; (3) only TT without RR; and (4) none. Accordingly, we propose four-group balance check rather than balance check within RR or TT as Dupas showed in Table 2 in her original paper, to enable us to compare among four subgroups.

		RR for 8th graders in randomly selected schools	
		1	0
TT for randomly selected schools	1	(TT, RR)	(TT, no RR)
	0	(no TT, RR)	(no TT, no RR)

Figure 2: Subgroup settings

Accordingly, four types of impact could be estimated: 1. the impact of both RR and TT: (1)-(4) 2. the impact of RR conditional on TT: (1)-(3) 3. the impact of RR without TT: (2)-(4) 4. the impact of TT without RR: (3)-(4)

The original article's strategy compares only the effect of RR against TT; however, this new empirical framework allows us to better estimate alternative intervention bundles, such as RR and TT, RR—TT. We will come back to this merit when we interpret the results.

Spillover Effect of RR: ITT and ToT

The treatment of RR is on HIV related information, and we should check whether this treatment spills over to the untreated cohorts. Spillover effect distinguishes the Intention to Treat (ITT) from Treatment to the Treated (ToT).

The difference between the estimated means in the follow equation is the Intention to Treat (ITT) estimate of RR.

Intention to Treat = mean Y of 8th graders in RR schools - mean Y of 8th graders in non-RR schools

This precisely answers the following question. What is the net impact of placing a group of students

under RR on their sexual behavior and teen pregnancy? This is exactly the question policy makers are interested in. Meanwhile, there is another question which prompts further investigation. What is the impact of learning HIV relative risk information on sexual behaviors and teen pregnancy? This impact is referred to as the effect of Treatment on the Treated (ToT). The treatment is offering HIV relative information to 8th graders in selected schools; however, the 'treated' are the set of students who are actually learning this set of information (in 8th grade and potentially in other grades). If we neglect the spillover of the treatment and assume the compliance rate is 100

Falsification Test

First of all, we make it explicit that two treatments are at different levels.

Treatment 1: TT for randomly selected schools Treatment 2: RR for 8th graders in randomly selected schools

The treated 1: TT schools The control 1: non-TT schools The treated 2: 8th graders in RR schools The control 2: 8th graders in non-RR schools

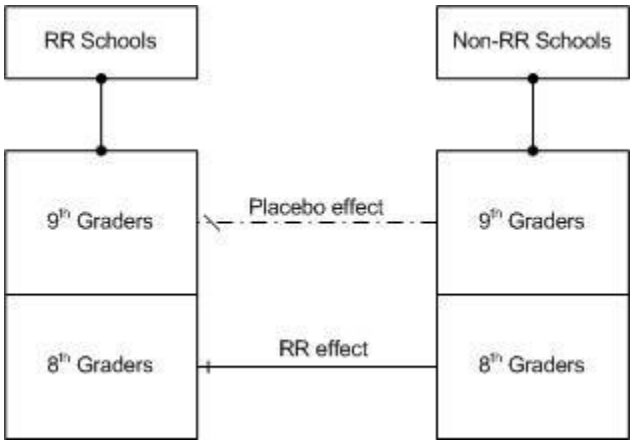

Figure 3: Settings for Falsification Test

We could detect whether there is spillover through the following falsification tests through using the 8th graders to detect the effect of RR and comparing the 9th graders for placebo treatment of RR. The examples below further illustrate how we will conduct falsification tests.

Regression 1: the Impact of RR without TT

Specification equation: ITT_{RR} = mean Y of 8th graders in RR schools - mean Y of 8th graders in non-RR schools Falsification equation: $ITT_{PLACEBO}$= mean Y of 9th graders in RR schools - mean Y of 9th graders in non-RR schools Given that only 8th graders are exactly treated by RR, the effect of RR on Y is falsified when the expected value of Y among 9th graders is significantly different between the treated (RR, no TT) and the control (no RR, no TT).

Regression 2: the Impact of RR—TT

Specification equation: $ITT_{RR|TT}$= mean Y of 8th graders in schools with RR and TT - mean Y of 8th graders in schools with TT Falsification equation: $ITT_{PLACEBO}$= mean Y of 9th graders in schools with RR and TT - mean Y of 9th graders in schools with TT Similarly, between the treated (RR, TT) and the control (no RT, TT), $ITT_{RR|TT}$ should only be detected among 8th graders other than 9th graders. Otherwise, significant difference of outcome among 7th graders in the treated and the control falsifies $ITT_{RR|TT}$.

Discussion of Findings

In this section we present the results of our analysis. In Part 1, we discuss the effect of RR, including the 'placebo test' of the control cohort. In Part 2, we present results for three other relevant subgroups: RR and TT, RR conditional on TT, and TT only.

We estimate the effect of RR on the probability that girls have started childbearing (Table 3 of the original article) and the self-reported sexual behavior of boys and girls who joined secondary school (Table 6 of the original article). These two tables focus on the main outcomes of interest we care about, both of which are imperfect but nonetheless defensible measures of 'risky' sex. [1]

[1]We are also concerned with the age gap between sexual partners. However, we believe that the model specification in the original paper addressing this question, Table 4: Age Gap between Girls Who Have Started Childbearing and their Partner, is biased due to the problem of 'censoring by death', and so we do not replicate it here. The problem is that some girls' lack of a post-treatment condition (childbearing) renders their values of the quantity of interest (age gap between girl and baby's father) undefined. Simply removing these girls from the analysis could bias the results in several different ways, as Dupas acknowledges in her discussion on selection issues (pp 18). Dupas makes the case that in order for the model to hold, we need to assume that the propensity to get pregnant (when engaging in unprotected sex) is independent of the age of the partner. She argues that there are several reasons this might not be true, but the bias could go either way, so we should compare the results for childbearing girls to those of girls who joined secondary school to assess the overall robustness of the findings. This does not seem like an adequate justification to us, and so we leave out the age gap for girls who started childbearing (Table 4 in the original paper), and take care to interpret the results of self-reported sexual behavior (Table 6 in the original paper) only for

As discussed in the original paper, we note that the students who join secondary school are not representative of all teenagers. They are likely to be wealthier and arguably smarter than those that do not. Also, in Kenya childbearing is not compatible with schooling, so the sexual behavior of girls who join secondary school may differ systematically from those that do not.

Verifying Balance between Subgroups In order to make valid comparisons across the series of subgroups we suggest are relevant for analysis, we first verify that the randomization holds for these subgroups on a range of school-level and individual-level characteristics. We confirmed balance between subgroups and report the results in the Appendix. Similar to the original article (see Table 2), we found that the only significant differences between groups were class size (lower on average in RR treatment schools), the number of students who repeat grade 8 (lower in RR), and the number who go on to secondary school (those from schools in TT treatment less likely). We stipulate to the original paper on these points.

We also checked balance for the study sample at follow-up in order to assess whether the sub-groups are still balanced at the secondary school level (see Appendix). We found no significant differences between the subgroups in this case. Most importantly, we found no significant differences for the percent in secondary school.

Part 1: The Effect of RR Dupas finds that providing information on the relative risk of HIV infection by partners age leads to a 28 percent decrease in teen pregnancy. She also finds that the likelihood that a girl's partner is more than 5 years their senior is significantly lower compared to the control group (for secondary school students).

the subset of students who choose to join secondary school, not for the whole student population.

Table1A—The Effect of RR on the Probability That Girls Have Started Childbearing

	Childbearing	Childbearing Unmarried	Childbearing Married
Effect of RR no TT			
Grade 8			
Coefficients	-0.0073	-0.00155	-0.005
Clustered Standard Errors	(0.00959)	(0.00593)	(0.0085)
Observations	3079	3079	3079
Mean(RRwithoutTT=0)	0.048	0.0168	0.0316
Grade 9			
Coefficients	-0.00577	0.01072	-0.0043
Clustered Standard Errors	(0.0194)	(0.0158)	(0.0156)
Observations	2548	2548	2548
Mean(RRwithoutTT=0)	0.124	0.061	0.083

Notes: The means reported in this table are the means for the dependent variables. All the regressions are controlled for individual characteristics and the primary school characteristics as in the original table.

Table1B—Effect of RR on Self-Reported Sexual Behavior for Students Who Joined Secondary School

	Has had sex with multiple partners	Currently has a regular partner	Partner >5 years older	Ever received money from current partner	Ever had sex	Ever had sex but never used a condom	Used a condom at last sexual intercourse
1,Girls							
Grade 8							
Coefficients	-0.010	0.085**	-0.076	-0.114	0.104**	0.056	-0.002
Clustered Standard Errors	(0.0224)	(0.0388)	(0.0626)	(0.1445)	(0.0452)	(0.0412)	(0.1100)
Observations	1514	1512	171	159	1517	1517	200
Mean	0.0309	0.1138	0.0753	0.6642	0.1888	0.1027	0.3882
Grade 9							
Coefficients	0.006	-0.046	-0.031	0.331**	0.023	0.012	-0.103
Clustered Standard Errors	(0.0210)	(0.0387)	(0.0546)	(0.1418)	(0.0576)	(0.0449)	(0.1426)
Observations	1325	1326	176	168	1335	1335	219
Mean	0.0415	0.1404	0.0727	0.6541	0.2139	0.1025	0.4523
2,Boys							
Grade 8							
Coefficients	0.136**	0.053	-0.019	0.229	0.111**	0.014	-0.001
Clustered Standard Errors	(0.0455)	(0.0340)	(0.0178)	(0.1324)	(0.0498)	(0.0552)	(0.0539)
Observations	1905	1869	224	247	1912	1912	798
Mean	0.2192	0.1367	0.0248	0.4529	0.5265	0.3384	0.2857
Grade 9							
Coefficients	0.047	0.003	-0.035	-0.128	0.013	0.061	-0.017
Clustered Standard Errors	(0.0474)	(0.0364)	(0.0303)	(0.1475)	(0.0501)	(0.0495)	(0.0579)
Observations	1390	1385	240	253	1395	1395	679
Mean	0.2873	0.1941	0.0404	0.4468	0.5888	0.2901	0.4200

Notes: The mean here is the mean for the dependent variable when setting RRwithoutTT=0.All the regressions are controlled for individual characteristics and secondary school characteristics as in the original table.

$**$ Significant at 5 percent level

$*$ Significant at 10 percent level.

In striking contrast to Dupas' results, we find that the RR program has an effect on the probability that girls have started childbearing (see Table 1A and Table1B for results). For girls that join secondary school, we also find no evidence that the age gap between girls and their partners has reduced. We find some consistency with the original results in tests of having a regular partner (75% more likely with RR) and ever had sex (55% more likely with RR) (percentage change calculated by dividing the coefficient by the mean of the control cohort at baseline).

The original article places most of the emphasis on effects for girls. Turning to boys, we find that RR increases the likelihood of ever having sex by 21% and having sex with multiple partners by 62%. Furthermore, there is no simultaneous rise in condom use, for boys or girls.

Given that the original article also mentions the possibility of increased anal sex (highly risky in terms of HIV transmission) and abortions as particular concerns when teenage girls switch from older to younger partners, and we find both increased sex and reduced incidence of pregnancy, there is real reason for concern that the RR program is negatively affecting both boys and girls.

Placebo Test: The Effect of RR on the Control Cohort Our alternative estimation strategy offers a falsification test for the problem of negative spillover between the cohorts. In the original paper, Dupas discussed the possibility that untreated girls in the treatment environment (the control cohort) may be more likely to engage in sex with older partners because the newly turned down 'sugar daddies' may try their luck with other girls instead (pp 22).

We estimated simple difference models for girls in the 9th grade (control cohort for the effect of RR on childbearing) and 9th grade (control cohort in secondary school, for the effect of RR on sexual behavior). We find no evidence of a negative spillover effect for the control cohort in either case (see Table 1A and 1B).

Part 2: The Effect of Other Subgroups The study design of this experiment makes it possible for us to compare the results of our findings in Part 1 to other situations in which the RR treatment was carried out along with the state TT program. In order to present the full set of information, we also include an assessment of the pure effect of TT.

Overall, the effect of RR and TT and the effect of RR conditional on TT both increase the likelihood of students reporting they have ever had sex (as did RR alone). However, we also see an increase in the likelihood of using a condom at last sexual intercourse, suggesting that both of these subgroups are preferable to RR alone. Finally, the

reduced age gap between partners that Dupas finds in the original paper seems entirely driven by the RR conditional on TT subgroup.

We discuss the findings for each subgroup below.

i. Effect of RR and TT Having both RR and TT programs (compared to none) has no effect on the probability that girls have started childbearing (See Table 2A and 2B).

Table2A—The Effect of RR and TT on the Probability That Girls Have Started Childbearing

	Childbearing	Childbearing Unmarried	Childbearing Married
Effect of RR and TT			
Grade 8			
Coefficients	-0.0166	-0.0107	-0.0063
Clustered Standard Errors	(0.0147)	(0.00665)	(0.0109)
Observations	2960	2960	2960
Mean(RRandTT=0)	0.048	0.0618	0.0316

Notes: The means reported in this table are the means for the dependent variables. All the regressions are controlled for individual characteristics and the primary school characteristics as in the original table. Both treated (grade 8) and untreated (grade 9) groups are exposed to TT, so there is no falsification test in this table.

For secondary school girls, RR and TT increases the likelihood of currently having a regular partner by 95%, increases the likelihood of ever receiving money from current partner by 31%, and increases the likelihood of using a condom at last intercourse by 51%.

For secondary school boys, RR and TT increases the likelihood of those who have ever had sex by 19%.

Table2B—Effect of RR and TT on Self-Reported Sexual Behavior for Students Who Joined Secondary School

	Has had sex with multiple partners	Currently has a regular partner	Partner >5 years older	Ever received money from current partner	Ever had sex	Ever had sex but never used a condom	Used a condom at last sexual intercourse
1,Girls							
Grade 8							
Coefficients	0.007	0.108**	-0.072	0.211**	0.065	-0.010	0.197**
Clustered Standard Errors	(0.0230)	(0.0377)	(0.0403)	(0.1000)	(0.0405)	(0.0361)	(0.0964)
Observations	1511	1505	172	163	1514	1514	201
Mean	0.0309	0.1138	0.0753	0.6642	0.1888	0.1027	0.3882
2,Boys							
Grade 8							
Coefficients	0.072	0.020	-0.033	0.169	0.101**	0.072	0.030
Clustered Standard Errors	(0.0462)	(0.0394)	(0.0213)	(0.1514)	(0.0475)	(0.0462)	(0.0516)
Observations	1924	1894	223	245	1933	1933	807
Mean(RRandTT=0)	0.2192	0.1367	0.0248	0.4529	0.5265	0.3384	0.2857

Notes: The mean here is the mean for the dependent variable when setting RRandTT=0.All the regressions are controlled for individual characteristics and secondary school characteristics as in the original table.Both treated (grade 8) and untreated (grade 9) groups are exposed to TT, so there is no falsification test in this table.

** Significant at 5 percent level

* Significant at 10 percent level.

ii. Effect of RR conditional on TT Having both RR and TT programs (compared to TT alone) decreases the likelihood that girls have started childbearing by 77% (see Table 3A and 3B). This seems surprising, but we believe it is related to the fact that the TT treatment alone actually increases the likelihood of childbearing (by 55%, as discussed in part iii). So, conditional on the existence of TT, RR is effective at reducing the likelihood of childbearing, whereas compared to no treatment, we see no effect for RR.

Table3A—The Effect of RR conditional on TT on the Probability That Girls Have Started Childbearing

	Childbearing	Childbearing Unmarried	Childbearing Married
Effect of RR conditional on TT			
Grade 8			
Coefficients	-0.0217	-0.020**	-0.003
Clustered Standard Errors	(0.014)	(0.00712)	(0.00993)
Observations	2909	2909	2909
Mean(RRconTT=0)	0.06	0.026	0.035
Grade 9			
Coefficients	-0.0167	0.0159	-0.0265
Clustered Standard Errors	(0.0176)	(0.0127)	(0.0167)
Observations	2432	2432	2432
Mean(RRconTT=0)	0.125	0.062	0.079

Notes: The means reported in this table are the means for the dependent variables. All the regressions are controlled for individual characteristics and the primary school characteristics as in the original table.

** Significant at 5 percent level

* Significant at 10 percent level.

We see a similar effect for secondary schools. For girls, RR conditional on TT increases the likelihood of currently having a partner by 96%, decreases the likelihood that the partner is more than five years older by 118%, increases the likelihood of ever had sex by 60%, and increases the likelihood of using a condom at last intercourse by 72%.

For secondary school boys, RR conditional on TT decreases the likelihood that the partner is more than five years older by 76% and increases the likelihood of ever had sex by 28%.

Table3B—Effect of RR conditional on TT on Self-Reported Sexual Behavior for Students Who Joined Secondary School

	Has had sex with multiple partners	Currently has a regular partner	Partner >5 years older	Ever received money from current partner	Ever had sex	Ever had sex but never used a condom	Used a condom at last sexual intercourse
1,Girls							
Grade 8							
Coefficients	0.025	0.121**	-0.075**	0.144	0.107**	0.017	0.235**
Clustered Standard Errors	(0.0188)	(0.0364)	(0.0330)	(0.1016)	(0.0382)	(0.0293)	(0.0983)
Observations	656	645	89	87	656	656	107
Mean	0.0262	0.1269	0.0635	0.6721	0.1776	0.1028	0.3289
Grade 9							
Coefficients	-0.034**	-0.021	0.008	0.042	-0.063	-0.044	0.232
Clustered Standard Errors	(0.0105)	(0.0373)	(0.0610)	(0.1196)	(0.0479)	(0.0308)	(0.1541)
Observations	547	537	82	74	550	550	100
Mean	0.0395	0.1647	0.0441	0.7258	0.2517	0.1132	0.4886
2,Boys							
Grade 8							
Coefficients	0.070	0.006	-0.045**	0.195	0.138**	0.080	0.046
Clustered Standard Errors	(0.0412)	(0.0366)	(0.0251)	(0.1611)	(0.0439)	(0.0483)	(0.0620)
Observations	763	772	88	103	766	766	318
Mean	0.2118	0.1356	0.0597	0.3951	0.4992	0.3169	0.3013
Grade 9							
Coefficients	0.020	-0.028	-0.073	-0.105	0.047	0.009	0.056
Clustered Standard Errors	(0.0550	(0.0413	(0.0398	(0.1554	(0.0388	(0.0444	(0.0579
Observations	599	591	90	93	599	599	294
Mean	0.29734	0.1714	0.0658	0.6203	0.6008	0.3136	0.4076

Notes:The mean here is the mean for the dependent variable when setting RRconTT=0.All the regressions are controlled for individual characteristics and secondary school characteristics as in the original table.

$**$ Significant at 5 percent level

$*$ Significant at 10 percent level.

iii. Effect of TT Finally, we can compare the effects of RR in different contexts to the effects of the Kenya state teacher training program (see Table 4A and 4B). As previously mentioned, TT increases the likelihood of childbearing by 55%. However, we find no evidence that TT has an effect on secondary school girls' or boys' self-reported sexual behavior.

Table4A—The Effect of TT on the Probability That Girls Have Started Childbearing

	Childbearing	Childbearing Unmarried	Childbearing Married
Grade 8			
Coefficients	0.00879	0.00922$*$	-0.00049
Clustered Standard Errors	(0.00734)	(0.00422)	(0.00579)
Observations	4819	4819	4819
Mean(TTwithoutRR=0)	0.048	0.0168	0.036

Notes: The means reported in this table are the means for the dependent variables. All the regressions are controlled for individual characteristics and the primary school characteristics as in the original table.Both treated (grade 8) and untreated (grade 9) groups are exposed to TT, so there is no falsification test in this table.

Table4B—Effect of TT on Self-Reported Sexual Behavior for Students Who Joined Secondary School

	Has had sex with multiple partners	Currently has a regular partner	Partner >5 years older	Ever received money from current partner	Ever had sex	Ever had sex but never used a condom	Used a condom at last sexual intercourse
1,Girls							
Grade 8							
Coefficients	-0.017	0.004	0.012	0.006	-0.025	-0.013	-0.047
Clustered Standard Errors	(0.0116)	(0.0235)	(0.0456)	(0.0895)	(0.0251)	(0.0182)	(0.0772)
Observations	1925	1916	209	198	1928	1928	246
Mean	0.0309	0.1138	0.0753	0.6642	0.1888	0.1027	0.3882
2,Boys							
2,Effect of TT no RR							
Grade 8							
Coefficients	0.002	0.011	0.026	-0.007	-0.028	-0.008	-0.024
Clustered Standard Errors	(0.0220)	(0.0205)	(0.0223)	(0.0751)	(0.0262)	(0.0254)	(0.0368)
Observations	2379	2346	269	304	2385	2385	967
Mean	0.2192	0.1367	0.0248	0.4529	0.5265	0.3384	0.2857

Notes:The mean here is the mean for the dependent variable when setting TTwithoutRR=0.All the regressions are controlled for individual characteristics and secondary school characteristics as in the original table.Both treated (grade 8) and untreated (grade 9) groups are exposed to TT, so there is no falsification test in this table.

∗∗ Significant at 5 percent level

∗ Significant at 10 percent level.

Conclusion and Implications

In this paper, we have shown that the estimates of the treatment effect of RR based on the original article's strategy and model specifications are biased. Our alternative estimation strategy allows us to address two major problems of the original design: randomization at multiple levels, and interactive effects with another existing treatment.

Our analysis shows that RR increases sexual intercourse, but it does not reduce the age gap between girls and their partners. Furthermore, it has no impact on childbearing, the objective proxy for risky sex used in the original paper. Subgroup analysis demonstrates that something slightly different goes on when the RR treatment is combined with the state teacher training program. We find that out of all the interventions considered, RR conditional on TT is the most effective compared with RR without TT, RR&TT, and TT without RR in terms of reducing teen pregnancy (again as a proxy for risky sex). However, we may believe that the reason RR conditional on TT is more effective in this case is because TT alone actually increases pregnancy/risky sex.

Overall, the evidence suggests that RR is more harmful than good. We are particularly concerned that the rise in incidence of sex across all subgroups except for TT alone is only matched with a rise in condom use when combined with the TT program. It is also worth noting that the rise in sexual behavior we find in self-reported sexual behavior is inconsistent with the lack of a rise in likelihood of childbearing. Either more students are using contraception (which the secondary school data does not support in all subgroups), more are having abortions, or more are having anal sex (with is very risky for HIV transmission). For all these reasons, we believe the RR program should not be implemented more broadly without further studies that address these issues.

We hypothesize that something may be going on in these schools which have the full set of risk information available. For example, the health clubs that many teachers set up as a result of the TT program could become places for students to share information about contraception after they learn about the relative risk information from the RR program. A girl who decides to substitute an older partner for a younger one due to the RR information, and is also concerned about getting pregnant more now than previously because her younger partner does not have the same level of financial resources, may find that others the health club are going through similar experiences and can share information. More generally, the health clubs may provide a safe place for students to discuss issues related to HIV - and so condom use may be a likely topic at some point.

Policy implications

The original article is highly focused on the potential benefits of HIV risk ratio information treatment for teenage

girls. They are considered the most vulnerable group in combating HIV/AIDS in Africa because they have higher HIV prevalence rates than men, and are also contracting HIV at higher rates. However, we also address the interest of boys in our reassessment of this article. Boys are also a vulnerable group because girls are three times more likely to have HIV, making any sex for them risky sex. In this paper we have shown that boys are worse off due to the HIV relative risk information treatment. According to our analysis, the RR program significantly increases sexual intercourse between teenagers, in which boys with lower HIV prevalence rate are exposed to increased risk of HIV/AIDS. The set of information about the benefit of boys is emphasized in our paper in order to shed lights on policy making.

There are still significant limitations to this f field experiment. We can only detect the marginal effect of an intervention, which is only partial equilibrium. However, for policy making, what we are really interested in is general equilibrium, in which we take into account all behavioral responses from all stakeholders. Given what we now know about how the treated group of the RR program (boys and girls) responds to this policy intervention, the next step in achieving a full understanding is an analytical framework to think through how the incentives of "sugar daddies" and untreated teenagers may change accordingly, and what the potential general equilibrium is if the RR program were to be scaled up.

Appendix

1, Balance Check

Table A—Verifying Balance between Groups in Terms of School Characteristics and Outcomes for Pre-program Cohort

	Subgroup Settings				Difference			
	RR	RR+TT	TT	Neither	RR	RR—TT	RR+TT	TT
	(1)	(2)	(3)	(4)	(1-4)	(2-3)	(2-4)	(3-4)
Panel A. School Characteristics at Baseline								
Class size	34.222	35.543	38.120	38.291	-4.069*	-3.577*	-3.748	-0.171
					(2.242)	(2.115)	(2.382)	(1.414)
Pupils' sex ratio	1.119	1.115	1.095	1.042	0.077	0.020	0.072	0.052
(girls/boys)					(0.090)	(0.112)	(0.102)	(0.062)
Teacher-pupil ratio	0.023	0.029	0.027	0.025	-0.003	0.002	0.004	0.0022
					(0.004)	(0.005)	(0.005)	(0.003)
Teachers' sex ratio	0.911	0.932	1.037	1.029	-0.118	-0.106	-0.097	0.008
(female/male)					(0.174)	(0.163)	(0.174)	(0.115)
KCPE results (2003)	251.333	247.457	249.470	252.488	-1.157	-1.990	-5.031	-3.041
					(5.430)	(5.481)	(5.464)	(3.680)
Panel B. Girls Enrolled in Eighth Grade in 2003: outcomes at the end of 2004								
Percent repeating class 8	0.209	0.208	0.247	0.244	-0.035	-0.039	-0.036	0.003
					(0.029)	(0.031)	(0.039)	(0.020)
Percent in secondary school	0.480	0.434	0.428	0.470	0.010	0.006	-0.036	-0.042*
					(0.038)	(0.035)	(0.039)	(0.024)
Percent in professional training	0.034	0.039	0.036	0.038	-0.004	0.003	0.001	-0.002
					(0.011)	(0.112)	(0.012)	(0.008)
Percent out of school	0.271	0.307	0.280	0.239	0.032	0.027	0.068 **	0.042**
					(0.029)	(0.032)	(0.032)	(0.020)
Percent married	0.066	0.010	0.081	0.073	0.007	0.019	0.027 *	0.008
					(0.014)	(0.020)	(0.016)	(0.011)
Percent who had begun	0.131	0.147	0.153	0.134	-0.003	-0.006	0.013	0.190
childbearing					(0.023)	(0.021)	(0.023)	(0.018)
Observations	613	599	2380	2403				
Panel C. Boys Enrolled in Eighth Grade in 2003: outcomes at the end of 2004								
Percent repeating class 8	0.220	0.220	0.223	0.229	-0.009	-0.004	-0.009	-0.005
					(0.032)	(0.032)	(0.032)	(0.021)
Percent in secondary school	0.533	0.484	0.514	0.527	0.006	-0.031	-0.043	-0.012
					(0.039)	(0.038)	(0.039)	(0.025)
Percent in professional training	0.011	0.007	0.011	0.019	-0.008	-0.003	-0.012	-0.008*
					(0.007)	(0.005)	(0.007)	(0.004)
Percent out of school	0.224	0.285	0.243	0.215	0.009	0.042	0.071 **	0.028
					(0.032)	(0.032)	(0.031)	(0.021)
Percent married	0.002	0.013	0.005	0.013	-0.011	0.08*	0.001	-0.007
					(0.008)	(0.004)	(0.009)	(0.005)
Observations	630	610	2385	2460				
Number of schools	36	25	125	127				

Notes: There are school averages. Panel A: school characteristics collected through school visits in 2004. Panels B and C: students' outcomes collected in 2004 for the cohort of students enrolled in grade 8 in 2003, which is the cohort just one year older than the cohort involved in the RR experiment. Data collected by asking whereabouts of students as their 2003 primary school. Standard deviations in brackets. Five schools did not have an eighth grade class in 2003 and there for are excluded for the table.

** Significant at 5 percent level

* Significant at 10 percent level.

Table B—Balance Check for Status of Study Sample at Follow-up

	Subgroup Settings				Difference			
	RR	RR+TT	TT	Neither	RR	RR—TT	RR+TT	TT
	(1)	(2)	(3)	(4)	(1-4)	(2-3)	(2-4)	(3-4)
Panel A. Girls								
Status unknown (attrited)	0.013	0.011	0.022	0.019	-0.003	-0.012	-0.008	0.003
					(0.007)	(0.028)	(0.006)	(0.009)
Percent repeating class 8	0.196	0.202	0.234	0.233	-0.037	0.001	-0.031	0.026
					(0.025)	(0.045)	(0.027)	(0.023)
Percent in secondary school	0.454	0.435	0.458	0.470	0.008	0.009	-0.034	-0.035
					(0.046)	(0.046)	(0.032)	(0.027)
Percent in professional training	0.063	0.052	0.045	0.046	0.009	-0.016	0.006	-0.006
					(0.017)	(0.015)	(0.011)	(0.009)
Percent out of school	0.298	0.317	0.269	0.260	-0.004	-0.001	0.057	0.009
					(0.044)	(0.049)	(0.029)	(0.025)
Percent dead	0.002	0.003	0.001	0.001	-0.002	-0.001	0.001	0.001
					(0.003)	(0.002)	(0.002)	(0.002)
Observations	879	715	3662	3897				
Number of schools	50	44	179	189				
Panel B. Boys								
Status unknown (attrited)	0.017	0.018	0.015	0.014	0.003	-0.012	0.004	0.001
					(0.005)	(0.009)	(0.005)	(0.004)
Percent repeating class 8	0.181	0.191	0.225	0.219	-0.038	-0.012	-0.028	0.003
					(0.023)	(0.035)	(0.024)	(0.019)
Percent in secondary school	0.515	0.509	0.477	0.485	0.030	-0.032	0.024	-0.004
					(0.030)	(0.048)	(0.031)	(0.026)
Percent in professional training	0.036	0.029	0.041	0.037	-0.001	-0.014	-0.008	0.010
					(0.009)	(0.016)	(0.010)	(0.008)
Percent out of school	0.273	0.277	0.262	0.265	0.008	0.052	0.012	-0.008
					(0.027)	(0.040)	(0.028)	(0.022)
Percent dead	0.001	0.001	0.001	0.001	-0.0002	0	-0.0002	-0.001
					(0.001)	(0)	(0.001)	(0.001)
Observations	944	845	3840	3864				
Number of schools	50	45	174	176				

Notes: Data collected by asking whereabouts of students at their 2004 primary school. Status as of July 2005. Study sample composed of students enrolled in eighth grade in 2004. The table presents school averages. Standard errors in parenthesis. Three schools did not have an eighth grade class in 2004 and there fore are excluded from the table.

** Significant at 5 percent level

* Significant at 10 percent level.