On the Doorstep of Adulthood: Economic Empowerment and Fertility Choices among Young Women in Tanzania*

Lars Ivar Oppedal Berge[a,b], Kjetil Bjorvatn[a], Fortunata Makene[c], Linda Helgesson Sekei[d], Vincent Somville[a,b], and Bertil Tungodden[b]

[a] *NHH Norwegian School of Economics*

[b] *Chr. Michelsen Institute*

[c] *Economic and Social Research Foundation*

[d] *NIRAS*

Young women leaving school have to make important choices about family and employment. But in a rural, low-income context, a lack of opportunities and decision-making power may severely limit the possibility of achieving

*We would like to thank FEMINA HIP for excellent cooperation throughout the research project, in addition we would like to thank Dr. Katanta Simwanza and Dr Goodluck Charles for their invaluable assistance in developing the training material. Special thanks to Juda Lyamai, Tausi Kida and Richard Ngilangwa who contributed to the success of the project, and to the numerous Master students of NHH who participated in the data collection. We have received valuable comments from XXXX and from seminar participants at XXxX. We have received financial support from the Research Council of Norway (226977, 250415 and 262675, 223269), CISMAC at the University of Bergen, The Royal Norwegian Embassy in Tanzania, Kavlifondet, NHH Norwegian School of Economics and Chr. Michelsen Institute. The project has been administered by The Choice Lab FAIR. check Dupas who shared her questionnaires

their life-long potential. We here report from a large-scale, long-term, cluster-randomized trial designed to study how female empowerment, in the health and economic domains, affects economic achievements and fertility decisions through changes in opportunities and decision-making power. The study is specifically designed to document potential complementarities between health and economic empowerment. Economic empowerment leads to a remarkable increase in self-employment and income, but also to an increase in fertility and teenage pregnancies. We attribute the increase in fertility to an income effect, where higher income makes it affordable to raise a child earlier. Health empowerment also leads to an increase in teenage pregnancies, initiated by a push towards more committed relationships. Our results carry an important message to policy makers about the potential, but also the complexity, of female empowerment interventions in low-income contexts.

We're smart enough to make these millions.
Strong enough to bear the children.
Then get back to business.

Beyonce
Run the World (Girls)

1. Introduction

Adolescence is a critical period in life. It's a time for decisions with potentially life-long consequences, on education, employment and sexual relations. Girls in low-income contexts are in a particularly vulnerable position, due to limited labor market opportunities and high risk of early pregnancy (Duflo, 2012; Field and Ambrus, 2008; Goldin and Katz, 2002; Rasul, 2008; Bailey, 2006; Miller, 2010). Empowering girls by increasing their economic opportunities and strengthening their decision-making power is therefore a major policy concern for governments, NGOs and donors, as reflected in the Sustainable Development Goals.

The linkages between economic empowerment and fertility decisions are however still poorly understood. Do women lack control over their fertility, leading to early childbearing and limited economic achievements (Goldin and Katz, 2002; Herrera et al., 2019; Rasul, 2008; Miller, 2010; Lundborg et al., 2017)? Or do limited economic opportunities push them to start a family early on (Heath and Mobarak, 2015; Jensen, 2012)?

We investigate these questions using a large-scale cluster randomized trial studying an empowerment program for adolescent girls in Tanzania, who at the time of the intervention were in the final year of secondary school. The program aimed at empowering the participants economically through a specific curriculum —"Build your life"— and in the health domain through another curriculum —"Protect your life"—.

In order to explore the linkages between economic empowerment and fertility decisions, we randomly assign the girls to one of four groups: a control group, the economic empowerment program, the health empowerment program, or both. We establish a rich data set

consisting of survey, experimental and medical data for the short term (a few weeks after the intervention), medium term (one year), and long term (three to four years), which allows us to study in detail the impact of the programs and underlying mechanisms.

We find that the economic empowerment program, alone or combined with the health program, causes a significant and enduring improvement in economic outcomes. Shortly after the training, the girls express greater business ambitions, and one year after the curriculum ended, we document an increase in their income generating activity. Three to four years later we observe that they still have significantly larger incomes than those who were not assigned to the economic empowerment program.

Contrary to our expectations, however, we do not find that any of the interventions reduced fertility. Quite the opposite, both programs lead to a marked *increase* in teenage pregnancy and, to a lower extent, in overall childbearing. We provide evidence consistent with the entrepreneurship program increasing early pregnancy through an income effect and the health program through an increase in committed partnerships.

To the best of our knowledge, our paper is the first to assess both separately and jointly programs aimed at economic and health empowerment. This is important both theoretically to understand the linkages between the two domains, as well as from a policy perspective, in order to explicitly compare the efficiency of different programs.

While the interventions differ, Duflo et al. (2015) takes a similar approach and investigates the role of education-subsidies alone or in combination with an abstinence curriculum in Kenya. They find that education subsidies reduce teenage pregnancy, while the abstinence curriculum has no impact. Another example comparing education and financial incentives is Buchmann et al. (2018) which investigates the impact of a broad-based education and empowerment program alone or in combination with financial incentives to delay marriage in Bangladesh. They conclude that the financial incentives reduce the likelihood of teenage childbearing, while the empowerment program has no impacts on childbearing and moreover that there are no interaction effects between the two programs. Finally, Ashraf et al (2020) study the impact of negotiation skills on educational outcomes in Zambia, focusing on eight-grade girls. In order to explore mechanisms, they compare the negotiation treatment with a more classical empowerment treatment as well as a

control group, and find that the negotiation treatment had a strong effect on enrollment in higher-level education in general and on higher-ability tracks in particular.

Bandiera et al. (2020) investigates the role of youth clubs in Uganda, offering a bundle of vocational training, information on sexual and reproductive health as well as a safe space to spend time. The authors find that this leads to a reduction in teenage pregnancies and an increase in business creation. Given that the program was offered as a bundle, it is however not clear which aspect was decisive in generating these positive results.[1]

Several studies consider either economic empowerment or health empowerment. For instance, there is a literature on how entrepreneurship training may improve economic opportunities (e.g. Berge et al., 2015; Karlan and Valdivia, 2011; Field et al., 2010). As pointed out by McKenzie and Woodruff (2014), impacts from these programs are most often modest and transitory. A plausible explanation of the large impact on business outcomes in our study is the fact that we target a population of young, unmarried women who presumably have fewer domestic obligations and hence more freedom to allocate their time and effort to income generating activities (Karlan and Valdivia, 2011; Berge et al., 2014).

Similarly, our paper relates to the literature on sexual and reproductive health training, which has been found to increase knowledge and change attitudes, but typically with less impact on behavioral changes (for an overview, see Berge et al., 2018). Important exceptions include Dupas (2011) and Dupas et al. (2018), who find that educating youth on HIV-risks reduces childbearing substantially, while the government's official abstinence-only HIV curriculum did not have any such effect.

The remainder of the paper is organized as follows. The next Section presents the research design, Section 3 the main results, the discussion of mechanisms, and provides additional evidence on the impact on fertility, while Section 4 concludes.

[1] Investigating the impacts from a similar intervention in Tanzania, Buehren et al. (2017) find no effects on childbearing or business startups, a result they ascribe to weaknesses in program implementation.

2. Research Design

In this section we describe the participants, explain the sampling and randomization, present the timeline of the study and describe the interventions, and finally present the empirical approach.

2.1. Participants and randomization

The end of adolescence and the transition into adulthood is a critical period of life that can define individual economic trajectories (Heckman and Mosso, 2014). We targeted girls in their last year of secondary school (Form IV), which for most would mark the end of education.[2] Hence, they would soon have to make important decisions about whether to start a family or to start a business or some other income generating activity, and we timed the empowerment programs precisely to be close to these decisions.

The participants were recruited from a sample of 80 public schools in rural and semi-urban areas of Tanzania (in the regions of Tabora, Singida, Morogoro and Dodoma). A map of Tanzania with the approximate location of the schools is shown in Appendix E. Boys-only schools were excluded from the sampling. stratified by region - size - remote? The schools were part of the network of our implementing partner, Femina Hip, a leading NGO with a mission of empowering the youth, implying that the schools received magazines and were kept up to date on relevant initiatives. All Form IV girls in these schools were invited to take part in the project, and we interviewed 3 483 girls at baseline.

The sample size and number of schools were powered to detect changes in pregnancy rate, which is the most demanding variable to measure and therefore serves as a conservative estimate for the other variables of interest.[3]

[2] In Tanzania, students must pass a national exam at the of Form IV in order to pursue their education. Students from rural public schools, and girls inn particular, are known to fail in large numbers. In our sample only 6 percent passed the exam.

[3] Taking into account the effect of clustering and the fact that we have three different treatment groups in addition to a control group, we have with the planned sample a power of 80 percent (with a five percent confidence level) to detect a decrease in pregnancy rate from 25 to 20 percent (using the approach of Hayes and Moulton (2012)).

After the baseline survey, we randomly allocated the 80 schools equally between the four treatment arms. That is 20 schools to entrepreneurship training, 20 schools to health training, 20 schools to receive both entrepreneurship and health training, and 20 schools to control. The randomization was blocked by school-size (below or above 40 girls in Form IV) and by region.

Table 1 shows that our randomization ensured similar treatment groups (MORE ON THIS); they are only significantly different on one baseline characteristic. B: Which one?

As evident from Table 1, the girls were aged 16–18 at baseline. They belong to households that typically do not own a business, reflecting that farming is the most common activity in these districts. Being poor and located in a rural setting, only 38 percent of the households have access to electricity, and 13 percent of the houses have roofs made of grass or mud. On average, the girls have meat for dinner twice a week (not reported in the table).

The table also shows that the participants were relatively less informed about business than health, as measured by the percentage score on a knowledge tests. This probably reflects the limited attention paid to entrepreneurship in the country's educational system.

When asked about their plans for the future (not reported in the table), only 14 percent report starting a business as their preferred choice, the most common plan being private sector employment, while very few wanted to work as a domestic or in farming. The least desired option by far was starting a family and staying at home.

In line with this, they report 22 years as the youngest suitable age for a woman getting married, and they themselves want to get married at the age of 25, and to have three children, the first one at age 26.

On sexual practices, 53 percent believe that girls their age have had sexual intercourse, and 54 percent agree to the statement "Girls in my age sometimes receive money or gifts for having sex with older men". Around 40 percent also report that girls their age often get sexually harassed.

Table 1: Baseline summary statistics by treatment arm.

	Control Mean (s.d.)	Health Mean (s.d.)	Economic Mean (s.d.)	Health & Econ Mean (s.d.)
Individual level				
Age*	17.565	17.583	17.486	17.614
	(1.033)	(.998)	(.947)	(.92)
Age > 17	.475	.504	.475	.516
	(.5)	(.5)	(.5)	(.5)
Cognition	.62	.558	.661	.652
	(.486)	(.497)	(.474)	(.476)
Risk averse	.482	.438	.455	.52
	(.5)	(.496)	(.498)	(.5)
Health knowledge	.563	.582	.557	.572
	(.215)	(.206)	(.214)	(.221)
Business knowledge	.449	.457	.458	.481
	(.257)	(.258)	(.264)	(.277)
Household level				
Wealthy household	.539	.593	.578	.518
	(.499)	(.492)	(.494)	(.5)
Household owns a business	.282	.244	.243	.226
	(.45)	(.43)	(.429)	(.418)
Woman headed household	.183	.198	.223	.191
	(.387)	(.399)	(.416)	(.394)
School level				
Remote school	.428	.462	.472	.411
	(.495)	(.499)	(.499)	(.492)
N girls	55.014	59.326	66.275	58.444
	(14.971)	(15.997)	(22.145)	(11.003)
Obs.	869	852	938	820

* Age is shown for information but is not included as a covariate in the estimations, as specified in the pre-analysis plans we instead use the binary variable "Age > 17". To test whether the covariates correlate with treatment assignment we compare each treatment separately to the control in OLS regressions where the covariates are the independent variables. At the 5 percent level of significance, we don't reject the hypothesis that all coefficients are equal to zero when we compare the control with the health group ($F(10,39)=1.99$, p-value=0.06), and the combined arm ($F(10,39)=1.3$, p-value=0.27) but we do in the comparison with the entrepreneurship group ($F(10,39)=4.16$, p-value<0.05).

When asked about whether they would be happy if they became pregnant next year, 80 percent said no whereas the remainder were not sure, and only two percent said yes. They were equally confident that their parents would not be happy about them becoming pregnant the following year (80 percent), but less confident about pregnancy being perceived as negative by the boyfriend (56 percent were unsure and 9 percent thought he would be happy) or society as a whole (30 percent were unsure about whether a child would bring about more respect, while 5 percent thought that it would).

2.2. Interventions

The aim of the interventions was to empower the girls by expanding their opportunities and strengthening their decision-making power on the economic and reproductive health domain. The economic empowerment program provided the girls with knowledge on how to establish and run their own business and sought to inspire them to do so. Topics included customer care, marketing, record keeping, pricing of products, personal finance, and sessions aiming at improving entrepreneurial mindset and self-confidence.

The aim of the reproductive health empowerment program was to enable the girls to take control of their own body and health, by providing information and guidance about contraception and the consequences of risky sexual behavior, as well as making the girls aware of gender equality rights.

In both treatments the girls received their own copy of either a "build your life" or a "protect your life" booklet, and in many cases this was one of very few school related books the girls had. Those who received both empowerment programs, received both booklets.[4]

In order to strengthen the external validity and the scalability of the programs, they were implemented by local teachers at the schools. Selection of the teachers was done by asking the girls to name two teachers they trusted and could talk with, and then the headmasters would make appointments based on these recommendations. The teachers then attended a one-week instructor session organized by Femina Hip (two weeks for the

[4]For a complete list of topics in the training program, see Appendix F.

teachers involved in the combined treatment).

The empowerment sessions were offered to all the Form IV girls at the treatment schools, and took place in a classroom setting during the after-school hours, which normally involve sports, games, homework, discussions on current issues, etc.

Both the economic and the health programs had 8 weekly sessions of 1.5 to 2 hours, 1 session per week, while those who got both treatments received 16 biweekly sessions. The control group girls carried on with their normal after school activities of sports and games.

The participants attended on average 6.88 sessions in the health treatment and 6.65 sessions in the economic treatment. In the combined treatment, they attended on average 7.08 health sessions and 7.04 economic sessions. In an evaluation of the training program, more than 98 percent either agreed or strongly agreed that it was very useful for them; that it provided them with new information; and that it was very well organized.

2.3. Surveys

The baseline survey was conducted in the spring of 2013, and the interventions were introduced in August and September 2013. The immediate impact of the treatments was evaluated in a short-term follow up survey and lab-in-the field experiment conducted in October 2013, a few weeks after the training programs ended, measuring changes in knowledge, behavior, views on gender equality and empowerment.

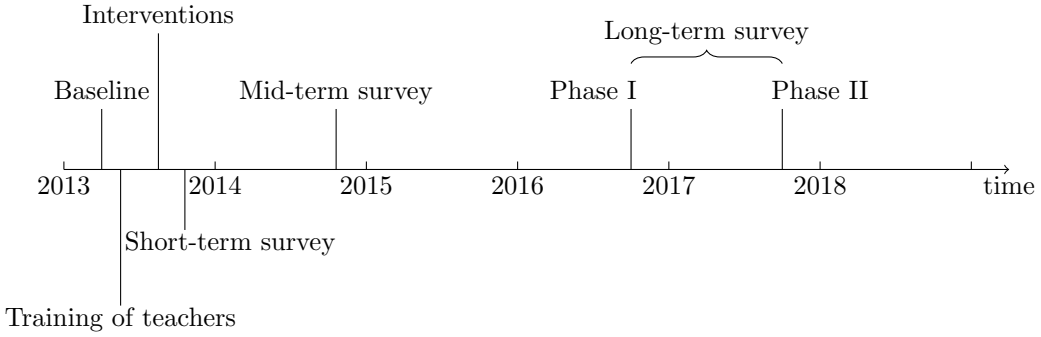


Figure 1: Timeline

Mid-term data were collected in September and October 2014. Since the large majority of the girls were expected to have quit school, we interviewed them by telephone. The aim of this survey was in particular to capture behavioral and well-being changes.

Finally, the long-term data collection started in June 2016. It was divided into two steps. First, we contacted the subjects by phone in order to learn their new place of residence and to administer a first interview on the phone. In the second phase, we sent teams of enumerators to meet the subjects in person for a face-to-face interview and to do pregnancy, syphilis and malaria tests. By the end of 2016, we had managed to interview 88 percent of the baseline sample on the phone and 72 percent in person. To reduce the attrition rate further, we organized a second phase in the Fall of 2017 to find and meet the participants that we missed in 2016. We interviewed 394 participants in this second phase, bringing the tracking rate in the face-to-face interviews up to 83 percent.

As we show in Appendix A, the rates of attrition in the different surveys don't differ significantly by treatment arm.

2.4. Empirical approach

The conceptional framework that we use to guide our empirical analysis is based on the hypothesis that choices are shaped by opportunities and decision-making power. Empowerment means expanding opportunities and increasing decision-making power, allowing people to make better choices for themselves, in our context both on the economic and on the health domain. Acknowledging the potential linkages between the two domains, we formulate the following three hypotheses:

First, we hypothesize that there is a *direct* effect of the economic program on economic outcomes and the health program on health outcomes.

Second, given the linkages between income and fertility, we hypothesize that there may be important *indirect* effects of the programs. For instance, economic empowerment may reduce fertility by making it more attractive to start a business rather than to start a family. Similarly, health empowerment may increase labor marked participation by

reducing exposure to sexually transmitted diseases and unwanted pregnancy.

Third, there may be *complementarities* between the programs. The effect of business training may be strengthened by health training, for instance through increased self-control or delayed pregnancy. And, similarly, the power over one's own sexuality may have limited effect if the participants lack feasible economic alternatives to starting a family.

In order to test these hypotheses, for each outcome Y_{ij} of individual i from school j we estimate the intention to treat estimators (ITT) using ordinary least squares:

$$Y_{ij} = \alpha + \beta_1 E_j + \beta_2 H_j + \beta_3 EH_j + \gamma X_{ij} + \epsilon_{ij} \tag{1}$$

Where E_j is equal to one if school j received the entrepreneurship training (only) and to zero otherwise, H_j indicates whether school j received the health training (only), and EH_j indicates whether school j receiving both treatments. X_{ij} is a vector of pre-specified covariates from the baseline defined at the individual or the school level. In appendix X we also report the estimates without those covariates.

According to our first hypothesis, on the direct effect of the programs, we expect that β_1 is positive for economic outcomes and that β_2 is positive for health outcomes. If β_1 is positive also for health outcomes, and β_2 is positive for economic outcomes, then this would be evidence in favor of the second hypothesis, namely that of an indirect treatment effect. Finally, if there are positive complementarities between the two programs, we would expect β_3 to be larger than $\beta_1 + \beta_2$.

We cluster the standard errors at the school-level. To adjust the p-values, we follow the procedure described in Benjamini and Hochberg (1995) to control for the false discovery rate and we group the tests by families.[5]

In our pre-analysis plan, we specified that we would test for heterogeneous impacts along four dimensions. We selected dimensions at the individual level —age and cognition— that could in principle directly influence the students' understanding of the program,

[5]We explain the procedure and the construction of the families in Appendix C.

fertility and economic choices. We also selected one dimension at the household level —wealth— and one at the school level —geographical remoteness—. Our thinking was that richer household may have different opportunities to set-up a business (e.g. better access to capital) or to delay marriages and pregnancies, and that more remote schools are located in environments that are less conducive to income generation and women's independence more generally. This analysis is presented in Appendix B. Overall, we don't observe significant heterogeneous impacts of the programs.

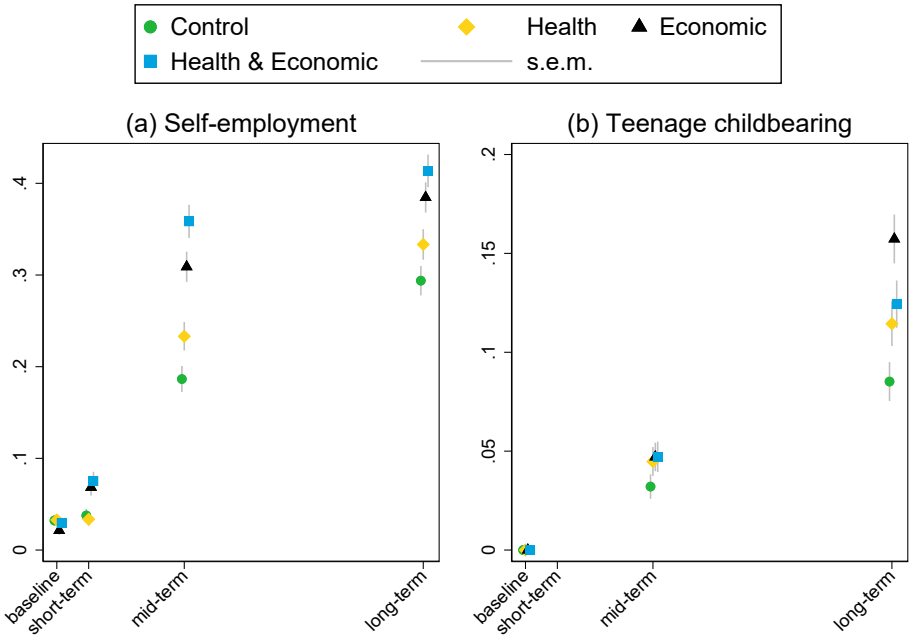
3. Results

We here present the results of our study, starting with the main results, that is, on self-employment and fertility, and then move on to mechanisms, in terms of opportunities and decision-making power. We close the section with a discussion of well-being, describing the impact of the programs on total income and happiness.

3.1. Main results

In Figure 2 we show the proportion of self-employed and of teenage pregnancies, by treatment arm and by survey. Very few of the participants had their own income generating activity at baseline and in the short-term when they were still in school. A year later, 19 percent of the girls from the control schools are self-employed. The proportion is slightly larger in the health group and it is around 30–35 percent among those assigned to the economic empowerment programs. In the long-term, the proportion of self-employed as increased to around 40 percent in the economic arms and lies around 30–35 percent among those who were not assigned to the economic program. Teenage pregnancies are also increasing in time, reaching almost 10 percent in the control group in the long-term, compared to around 15 percent in the economic arm and around 13 percent in the helth and the combined health economic arms.

Table 2 shows the estimated impacts on business development. Compared to Figure 2, the Table includes pre-specified controls in the estimation, and takes clustering into account



Note: Figure (a) shows the proportions of self-employed and Figure (b) shows the proportion of teenage pregnancies.

Figure 2: Self-employement & teenage pregnancies.

in calculating standard errors, as well as multiple hypothesis testing when calculating p-values. We observe that the economic empowerment program has had a remarkable effect on business creation and revenues. Columns (1) to (3) demonstrate that business training leads to a near doubling of the share of participants reporting having an income generating activity in the short and mid term (2 months and 1 year), and a one-third increase the long term (3-4 years). Typical business activities reported are selling street food, fruits and vegetables, charcoal or clothes, or braiding hair. The impact on business startups is consistent across the two groups that received this kind of training, that is, the economic empowerment group and the combined treatment group.

As a result of the increased business activity, we also observe a significant increase in business incomes, columns (4) to (7). The long-term evidence shows average weekly incomes (after winsorizing at the top 1 percent) of 9087 Tsh (PPP USD 10) for the economic empowerment group; 9093 Tsh (PPP USD 10) for the combined treatment group; 8377 Tsh (PPP USD 9) for the health group; and of 6654 Tsh (PPP USD 8) for the control group, hence a 25 percent increase in income from economic empowerment. In sum, these findings strongly support our hypothesis of a direct effect of the economic

empowerment program on our key economic outcome.

The table also shows that the health intervention doesn't have any significant impact on the economic outcomes, and we observe that the treatment effects from the health intervention are significantly different from those involving entrepreneurship training. Hence, we can refute the second hypothesis, on indirect effects, when it comes to the health intervention.

The fact that the outcomes on business startups and business income are very similar when entrepreneurship training is given alone or in combination with the health treatment implies that we can reject the hypothesis that there are complementarities between the two programs on economic outcomes. This suggests that economic empowerment can be achieved without empowerment on the sexual and health domain. We will have more to say about this later.

Table 3 shows results on the second main outcome, namely fertility. We find that teenage pregnancy, reported in columns (1) and (2), has increased markedly in all treatment groups, as reported in the long-term survey, column (2). Remarkably, the strongest treatment effect on teenage pregnancy comes from the economic empowerment program, where the long-term survey shows a near doubling compared to control. The lack of treatment effects on this dimension in the mid-term, at which point the median age was 18, probably reflects that the girls had just finished school, and with many in fact still in the school system, repeating classes to retake exams.

Total fertility is shown in columns (3) and (4), and we see that the evidence from the long-term survey are very much in line with those on teenage pregnancy, although somewhat weaker, which is probably due to a catching-up effect. In the long run, with the median age of the girls then being 21 years old, 32.6 percent of the control group members have started childbearing.

In sum, we find evidence a direct effect of the health intervention on fertility, as well as an indirect effect on fertility from the economic empowerment arm. The combined treatment has a very similar effect on fertility as the single treatments, which allows us to rule out any positive complementarity on this dimension.

The fact that both types of empowerment lead to an increase in early fertility is perhaps somewhat surprising, and in the next section we explore possible mechanisms.

Tables 2 and 3 provide the results of hypothesis tests using both unadjusted p-values (using the asterix $*$ symbol) and p-values adjusted for multiple hypothesis testing (using the star $\star$ symbol).

Table 2: Impacts on business development.

	(1)	(2)	(3)	(4)	(5)	(6)
	Self-employment			Business income (ihst)		
	short-term	mid-term	long-term	short-term	mid-term	long-term
Health	-.003	.038	.036	-.01	.284	.24
	(.009)	(.025)	(.032)	(.086)	(.201)	(.298)
Economic	$.033^{**}_{\star\star}$	$.115^{***}_{\star\star\star}$	$.093^{**}_{\star\star}$	$.288^{**}_{\star\star}$	$.721^{***}_{\star\star}$	$.796^{**}_{\star\star}$
	(.014)	(.032)	(.036)	(.116)	(.263)	(.305)
Health & Econ.	$.039^{**}_{\star\star}$	$.167^{***}_{\star\star\star}$	$.118^{***}_{\star\star\star}$	$.36^{**}_{\star\star}$	$1.157^{***}_{\star\star\star}$	$.83^{**}_{\star\star}$
	(.017)	(.038)	(.032)	(.155)	(.326)	(.323)
Mean Control	.038	.187	.294	.344	1.162	2.359
Obs.	2895	2994	3249	2895	2994	3249
Tests of equality of coefficients:						
Econ. - Health	$.035^{**}_{\star\star}$	$.077^{**}_{\star\star}$	$.057^{*}$	$.299^{**}_{\star\star}$	$.436^{*}$	$.556^{**}_{\star}$
	(.014)	(.03)	(.031)	(.12)	(.261)	(.251)
Econ. - Health & Econ.	-.006	-.052	-.025	-.072	-.437	-.034
	(.02)	(.042)	(.033)	(.174)	(.385)	(.298)
Health - Health & Econ.	$-.041^{**}_{\star}$	$-.129^{***}_{\star\star}$	$-.083^{***}_{\star\star}$	$-.371^{**}_{\star}$	$-.873^{**}_{\star\star}$	$-.59^{**}_{\star}$
	(.018)	(.037)	(.028)	(.163)	(.335)	(.288)
Health + Econ.	-.008	-.014	.01	-.083	-.152	.206
- Health & Econ.	(.022)	(.049)	(.047)	(.198)	(.443)	(.432)

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ."), tests of equality of impacts between arms (" Econ. - Health", "Econ. - Health & Econ." and "Health - Health & Econ."), and a test of complementarity between the Health and Economic empowerment programs ("Health + Econ.- Health & Econ."). The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 3: Impacts on fertility.

	(1)	(2)	(3)	(4)
	Teenage pregnancy		Started childbearing	
	mid-term	long-term	mid-term	long-term
Health	.011	.033**	.008	.051*
	(.011)	(.016)	(.017)	(.027)
Economic	.014	.08$^{***}_{***}$	.008	.056*
	(.01)	(.017)	(.014)	(.028)
Health & Econ.	.015	.045$^{***}_{**}$	.021	.025
	(.01)	(.016)	(.017)	(.03)
Mean Control	.032	.085	.056	.326
Obs.	3249	3249	2993	3262
Tests of equality of coefficients:				
Econ. - Health	.003	.047**	-.001	.005
	(.013)	(.019)	(.015)	(.022)
Econ. - Health & Econ.	-.001	.034*	-.013	.031
	(.012)	(.02)	(.016)	(.027)
SRH - Combined	-.004	-.013	-.012	.026
	(.013)	(.019)	(.018)	(.026)
Health + Econ. - Health & Econ.	.011	.067$^{**}_{**}$	-.005	.082$^{**}_{*}$
	(.017)	(.026)	(.023)	(.039)

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ."), tests of equality of impacts between arms (" Econ. - Health", "Econ. - Health & Econ." and "Health - Health & Econ."), and a test of complementarity between the Health and Economic empowerment programs ("Health + Econ.- Health & Econ."). The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ when the p-values are not corrected and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

3.2. Mechanisms

According to our conceptual framework, the empowerment program affects behaviors and outcomes by changing opportunities and decision-making power, and indirectly, through linkages between economic empowerment and fertility. We use this framework, illustrated by Figure 3, to discuss mechanisms, and do so sequentially, looking first at the program effects on *opportunities* and *decision-making power* and then on occupational choices and relationships. Finally we estimate the correlates of fertility, highlighting its link to income and stable relationships.

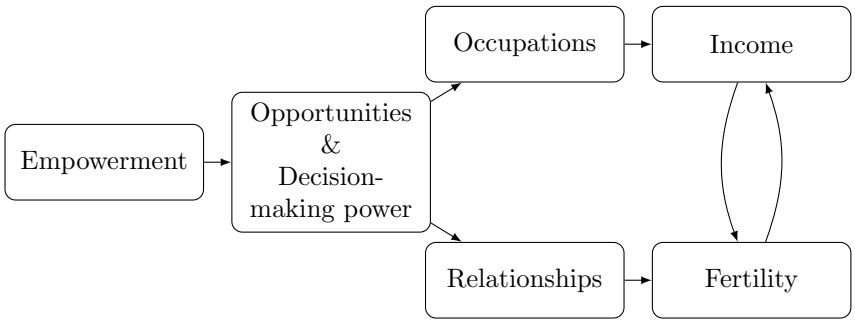


Figure 3: Schematic representation of the mechanisms.

We will document that the economic empowerment program expanded economic opportunities by increasing knowledge about entrepreneurship, and led to increased decision-making power by changing attitudes to gender norms and making the participants more willing to take risk (which may be important in order to start a business).

We also document that the health empowerment program led to increased decision-making power by strengthening self-control, by changing attitudes to gender norms, and by increasing the prevalence of stable relationships.

Finally, our correlation analysis suggests that fertility is positively associated with income and stable relationships. Digging deeper into the mechanisms, we also find that girls with higher income and those living in stable relationships are also more positive to having their first child in the near future, which indicates that these effects are based on the girls' own preferences rather those of the partner, parents or society.

3.2.1. Opportunities and decision-making power

The empowerment program may have expanded opportunities by increasing knowledge and awareness and strengthened decision-making power by fostering a stronger sense of control and by challenging gender stereotypes. The following quote from an essay written as part of the short-term survey, where the girls should envision themselves five to ten years into the future, shows an example of how the economic empowerment program appears to have affected both opportunities (awareness of a market for beauty products) and decision making power (the ability to follow up on intentions):

"*I am very grateful because I did not know that I as a woman have the ability to do what I had intended to. I would like to open a beauty products shop here in Mtamba because there are girls and women who like to beautify themselves but they cannot get access to beauty products. It is for this reason that I if am not successful in continuing with my education then it will be better for me to be an entrepreneur who sells various beauty products.*"

We measure knowledge through incentivized multiple choice questions on issues related to economics and health. For instance, one of the economics questions was about good customer service, where the correct answer is to try not to keep the customer waiting, and where one of the alternatives was to always promise a customer something, even if one is not sure about the ability to get it for them. An example of a health question concerned pregnancy at very young age, where the correct answer was that it was risky because the body is not fully developed, and where an alternative was that it proves that you are a grown-up woman. In Table 4, Panel A, we observe that the participants got around 68 percent of the health questions correct, and that there are no treatment effects on this dimension. Entrepreneurship was a more novel topic for the students, not covered by the school curriculum, and the control group score here was only 38 percent. We observe a positive treatment effect from the economic program on the economics questions, and we take this as evidence of expanded economic opportunities.

We capture decision-making power by several factors related to the girls' self control, their views on gender equality, and, on the economic domain, their *business mindset*.

We measure the participants' sense of being in control over their lives using an index

based on seven different questions, which we call *self control*. The questions were selected from the *psychological coping resources* measure of Pearlin and Schooler (1978): *I have little control about things that happen to me; I often feel helpless dealing with the problems of life; There is not much I can do to change important things in my life; On the whole, I am satisfied with myself; I am quite sure of myself; I certainly feel useless at times; I have a positive attitude towards myself.*

The economic dimension of gender empowerment is measured as the degree to which the respondents find it acceptable that the wife earns more money than the husband. On the health dimension of gender equality, we used the Demographic and Health Survey module on wife beating, and our measure here is the number of situations where the respondent finds such behavior unacceptable (0-5) references who used this and relate it to decision-making power.

We define business mindset as attitudes to risk and competition. We know from the literature that willingness to take risk and to enter into competition are important for economic success, and, moreover, that there are clear gender differences on these dimensions, potentially due to gender norms (Lars Ivar Oppedal Berge, Kjetil Bjorvatn, Armando Jose Garcia Pires, Bertil Tungodden Competitive in the lab, successful in the field?, Journal of Economic Behavior Organization Volume 118, October 2015, Pages 303-317). Our risk measure is based on a question about the general willingness to take risks in life (Dohmen et al., 2011), while willingness to compete is measured as the choice of competing rather than receiving a fixed payment in an incentivized adding numbers exercise (Niederle and Vesterlund).

Panel B of Table 4) shows that the health program, and in particular combined with the economic program, leads to a marked strengthening in self control. While the impact of the economic program alone goes in the same direction, its coefficient is not statistically significantly different from zero.

The economic program has a significant, positive impact on economic gender equality. The health program points in the same direction, although the effect here is smaller. We observe that health program, alone or combined, also leads to a reduction in the

acceptance of wife beating, but with no effect from economic program.[6]

Regarding business mindset, we observe a significant positive effect from the entrepreneurship program, alone or combined, on the willingness to take risks while there is no impact from the health program. Willingness to take calculated risks was an important message from the economic empowerment program, describing the entrepreneurial mind. We find no overall treatment effects on the willingness to compete.[7]

[6]These short-term effects are consistent with the findings of Dhar et al. (2020), in our case they however dissipate with time and disappear in the long-term.

[7]We do, however, observe interesting effects on the willingness to compete for certain subgroups, and refer to Almå s et al. (2020) for an analysis of this.

Table 4: Impacts on opportunities & decision-making power.

	Panel A: Opportunities		**Panel B: Decision-Making Power**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Health	Economic	Self-control	Gender equality		Business mindset	
			index	Economic	Health	Risk	Competitiveness
Health	.004	0	$.126^{***}_{**}$	$.136^{*}$	$.279^{***}_{***}$	-.01	-.027
	(.018)	(.012)	(.047)	(.077)	(.097)	(.246)	(.057)
Economic	.003	$.033^{***}_{**}$	.057	$.297^{***}_{***}$	-.089	1.059^{***}_{***}	.063
	(.017)	(.011)	(.041)	(.079)	(.07)	(.227)	(.048)
Health & Econ.	.016	$.062^{***}_{***}$	$.181^{***}_{***}$	$.302^{***}_{***}$	$.314^{***}_{***}$	$.857^{***}_{***}$	.01
	(.017)	(.014)	(.051)	(.076)	(.088)	(.225)	(.044)
Mean Control	.684	.38	-2.477	3.817	3.578	7.755	.332
Obs.	2898	2898	2891	2895	2890	2895	2912

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ."). The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ when the p-values are not corrected and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

3.2.2. Behavior

Given the changes that we see in opportunities and decision-making power in the short-term, we can expect that the programs affect economic and health behavior in the medium and longer run.

Table 5, Panel A, shows economic behavior, starting with occupational aspirations as captured by the short-term survey, and then moving on to their realized occupational choice at end-line. The economic program, alone or combined, has led to a substantial increase in the share of participants who want to become entrepreneurs, and who later indeed become one: a ten percentage point increase from a control group average of 29 percent. This increase comes at the expense of wage employment and unpaid domestic chores.

As we will see in XX the move into self-employment was a profitable one. We note that control group average income from self-employment is around fifty percent higher than for wage employment, and we observe that the economic empowerment program has led to a sharp increase in total income. A closer look at the data suggests that the higher income from self-employment comes mainly from more participants moving into self-employment (extensive margin) rather than higher income from those who are self-employed (intensive margin). Hence, the main impact of the entrepreneurship training appears to have been to inspire participants to start a business rather than teaching them how to run them. This seems reasonable, given the generic nature of the training program. We also note that earnings are larger among the self-employed compared to the wage workers on average, but the standard deviation is also larger, reflecting the riskiness of entrepreneurial activities. This shift towards higher returns and higher risk occupations is consistent with the increase in willingness to take risks documented in Table 4.

Panel B of Table 5 considers health behavior. We observe that the girls express a desire to marry at the age of 25 and have their first child at the age of 26, which is consistent with what we found in the baseline, six months earlier, and we find no treatment effect on these aspirations. We will return to how perceptions about fertility have changed when we visit them again at endline, two to three years later. Importantly, the health program, alone or combined, leads to a significant increase in stable relationships in the mid term,

that is, one year after the interventions.[8] From a control average of 28.6 percent, the health program increases the share of girls in stable relationships with 25–30 percent. There is no such effect from economic program. The increase in stable relationships is in accordance with the health program, where the "Protect your life"booklet stresses that *"Having a relationship with someone who is in multiple relationships is risky for your health and your well being. Healthy relationships are between two people who are honest and straightforward with each other."* The fact that we do find a mid-term effect on stable relationships, while the effect on childbearing only happens later, as evidenced from Table 3, suggests that the relationship causes the pregnancy, rather than the other way around. This mechanism is also remarkably consistent with the theory and the evidence from Kenyan schools reported by Duflo et al. (2015). In the long term, 75.4 percent of the girls are in a stable relationship, and the positive treatment effect is by and large gone.

On the other behaviors related to sexual and reproductive health, we do not observe any strong effects of the empowerment program. We do not observe any impact of the interventions on the number of sexual partners. There is a tendency that health training has increased condom use, and that the economic empowerment program has reduced exposure to unwanted sex, which is encouraging, given that self-employment could potentially have put them in a more vulnerable position.

[8]Stable relationships encompass being married or engaged, or being in an otherwise committed relationship.

Table 5: Impacts on economic behavior & health behavior

Panel A: Economic behavior

	Short-term: Plans		Long-term: Occupations				
	(1) Business	(2) Education	(3) Self-employed	(4) Wage	(5) Family business	(6) Student	(7) Domestic chores
Health	.022	-.023	.036	-.023	.019	-.04	.008
	(.028)	(.053)	(.032)	(.018)	(.017)	(.026)	(.028)
Economic	$.378^{***}_{\star\star\star}$	.045	$.093^{**}_{\star\star}$	-.026	$.031^{**}$	-.035	$-.063^{**}_{\star}$
	(.039)	(.052)	(.036)	(.02)	(.015)	(.028)	(.028)
Health & Econ.	$.432^{***}_{\star\star\star}$	.039	$.118^{***}_{\star\star}$	$-.046^{**}_{\star\star}$	.005	-.026	$-.051^{*}$
	(.046)	(.048)	(.032)	(.018)	(.016)	(.031)	(.028)
Mean Control	.151	.694	.294	.11	.07	.191	.335
Obs.	2894	2892	3249	3249	3249	3249	3249

Panel B: Health behavior

	Short-term: Plans		Stable relationship		Number of	Condom	Unwanted
	marriage	1st child	mid-term	long-term	partners	use	sex
Health	.161	-.002	$.085^{**}$	$.043^{*}$	.155	$.049^{*}$	-.007
	(.217)	(.23)	(.039)	(.022)	(.13)	(.025)	(.015)
Economic	.15	.19	-.024	-.015	-.115	.009	$-.038^{**}$
	(.248)	(.245)	(.029)	(.023)	(.11)	(.028)	(.016)
Health & Econ.	.085	.03	$.072^{**}$	.001	.016	.05	-.022
	(.226)	(.221)	(.034)	(.023)	(.113)	(.033)	(.017)
Mean Control	25.584	26.737	.286	.754	1.8	.815	.101
Obs.	2895	2863	2722	2895	2895	2506	2506

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ."). The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

3.2.3. Correlates of fertility

In this section, we shed further light on plausible mechanisms explaining the observed increase in early pregnancy by studying the correlates of fertility. We have observed that there is a causal effect of the economic empowerment program on income, and a causal effect of the health program on stable relationships. Our hypothesis is that both of these effects then lay the ground for the observed increase in fertility. Observing a positive correlation between these variables in the data would lend support to this hypothesis.

Indeed, this is what we find. Table 6 studies the correlates of fertility, where columns (1) to (3) consider only the control group, and show a strong positive and statistically significant association between income and childbearing and having a stable relationship and childbearing, and that these associations are robust the inclusion of the controls used in the previous tables. This indicates in particular that on average the substitution effect brought by better economic opportunities seems to be dominated by the income effect, since the coefficient of income would otherwise be negative.

In columns (4) and (5) we use the full sample and add indicators for the treatments. We note that the economic and health empowerment coefficients are reduced and lose their statistical significance once we introduce income and stable relationship in the model. We interpret columns (4) and (5) as additional evidence that the treatments have an impact through the income effect and through the stability of relationships.

This interpretation is consistent with the model proposed by Duflo et al. (2015) in which pregnancy is valued, and particularly so in stable relationships and when it's not too costly economically. More generally, standard models of fertility decision (Becker, 1960; Hotz et al., 1997; Schultz, 1997) emphasize that improving economic opportunities has two effects going in opposite directions: better opportunities increase the opportunity cost of having children, lowering fertility, but higher incomes also mean that people can afford to have more children (or have them earlier) and want to have more children (if children are a "normal good"). It is theoretically unclear which of the two is likely to dominate (Galor and Weil, 2000; Jones et al., 2010; Mookherjee et al., 2012).

Our results also harmonize with Lagerlöf (2015), who shows that in 18th and 19th century

Table 6: Higher income and stable relationships associate with higher fertility.

	(1)	(2)	(3)	(4)	(5)
Income	0.017***		0.011**		0.013***
	(0.004)		(0.005)		(0.002)
Stable relationship		0.156***	0.153***		0.145***
		(0.035)	(0.033)		(0.018)
Health				0.051*	0.026
				(0.027)	(0.028)
Entrepreneurship				0.056**	0.025
				(0.028)	(0.029)
Health & Entrepreneurship				0.024	0.001
				(0.030)	(0.028)
Remote school			0.009	0.043**	0.014
			(0.047)	(0.020)	(0.020)
Wealthy household			-0.070	-0.059***	-0.060***
			(0.048)	(0.018)	(0.020)
Cognition			-0.020	-0.051**	-0.024
			(0.053)	(0.020)	(0.022)
Age > 17			0.056	0.073***	0.043**
			(0.034)	(0.016)	(0.018)
N Form IV girls			0.000	-0.001	-0.001
			(0.001)	(0.001)	(0.001)
Woman headed hh.			0.031	0.055**	0.057***
			(0.053)	(0.022)	(0.021)
Business owner			0.008	0.007	0.005
			(0.031)	(0.020)	(0.021)
Constant	0.280***	0.249***	0.218**	0.349***	0.261***
	(0.024)	(0.026)	(0.090)	(0.043)	(0.046)
Observations	810	663	663	3249	2604
R-sq.	0.03	0.02	0.05	0.02	0.06

The table shows OLS estimates of the correlation between *income*, *stable relationships* and *childbearing*. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$. All the estimations include the covariates listed in 1.

Sweden, a good harvest one year led to higher marriage and birth rates the following year. Further evidence is provided by Ashraf and Galor (2011) who argue that in the period 1-1500 CE technological gains and increased land productivity increased population density.

The positive income effect does, however, contrast with the negative impact of employment on fertility documented by Heath and Mobarak (2015) and Jensen (2012). But note that these papers consider jobs in factories and call centers, which are not easily combined with raising children. In contrast, increased income in our setting comes from self-employment, with work closer to home and with more flexible work hours, which makes it easier to combine work with taking care of a child.

Fertility decisions however usually involve more than one decision maker, and there is an increasing interest among researchers in better understanding how couples make these decisions in low-income countries (e.g. Doepke and Tertilt, 2018; Doepke and Kindermann, 2019). Our data set is very rich but unfortunately lacks information about the partners and about how couples are taking decisions. We can nonetheless provide some descriptive evidence about the extent to which having a child is in line with the women's individual preferences, and her perception of her family's preferences and those of society.

At baseline and at endline, we asked the respondents: "Would you be happy to have a child next year?"; "Would your parents be happy if you have a child next year?"; and "Will people treat you with more respect if you have a child next year?". At baseline, only two percent stated that they would be happy to have a child, two percent thought their parents would be happy, and five percent thought they would receive more respect from society if they were to have a child at that point in their lives. A few years later, the picture is very different. Considering only those who do not have a child yet, 44 percent would be happy to get one, 37 percent think their parents would be happy and 45 percent say they would receive more respect.[9]

The answers to these questions strongly correlate with their incomes and with being in

[9]They answered on a five point scale from "strongly disagree" to "strongly agree". We transformed the answers in a binary variable equal to 1 if they strongly agree, agree or neither agree nor disagree. In the group that has a child already 24 percent would be happy to get another child next year, 26 percent think their parents would be happy and 32 percent say they would receive more respect. Note that we also asked those questions in the short-term survey, the answers are very similar to the baseline.

a stable relationship. Among those who have a stable relationship and earn their own income, 54 percent would be happy to have a child in the coming year. The proportion is 48 percent if they have a stable relationship but no income, 40 percent if they have an income but not a stable relationship and 29 percent if they have no income and no stable relationship. These numbers suggest that the increase in fertility caused by the empowerment programs were in line with the girls' individual preferences.

Table 7 shows the correlations between having an income, stable relationships, and whether the respondent would be happy to have a child next year, whether her parents would be happy and whether she would receive more respect. Column (1), (3) and (5) use the full sample and include a control for having a child already. The other columns only use the sample of respondents who don't have a child.

In line with the income effect discussed above, the table indicates that having a child becomes increasingly desirable when one has her own income. If there was a strong substitution effect, the coefficient of "Has own income" would be negative. That "stable relationship" correlates positively with "being happy to get a child" also indicates that the respondents have a preference for having children.

The Table also shows that the perceived expectations from the family and from society more broadly are aligned with the individual willingness to have a child: when they have their own income and a committed partner, they are more likely to be happy to get a child, more likely to think that their parents would also be happy and that people would treat them with more respect. In this sense, there is no evidence of a conflict between their own fertility preferences and the perceived preferences of others about their own fertility.[10]

.

[10]It is of course plausible that in the evaluation of their own happiness they integrated the preferences of their family and of society.

Table 7: Income and stable relationships correlate with willingness to have a child.

	(1)	(2)	(3)	(4)	(5)	(6)
	You would be happy		Parents would be happy		Would get respect	
Has own income	0.038*	0.060**	0.052***	0.085***	0.031	0.061**
	(0.021)	(0.027)	(0.02)	(0.028)	(0.02)	(0.027)
Stable relationship	0.145***	0.169***	0.161***	0.182***	0.169***	0.184***
	(0.024)	(0.026)	(0.02)	(0.024)	(0.021)	(0.024)
Has a child	-0.238***		-0.164***		-0.180***	
	(0.024)		(0.02)		(0.026)	
Health	-0.033	0.002	-0.024	-0.007	-0.02	-0.029
	(0.028)	(0.038)	(0.03)	(0.039)	(0.029)	(0.036)
Econ.	-0.039	-0.018	-0.012	-0.033	-0.013	-0.056
	(0.028)	(0.037)	(0.027)	(0.033)	(0.028)	(0.035)
Health & Econ.	-0.039	-0.036	-0.026	-0.023	-0.017	-0.072**
	(0.03)	(0.035)	(0.034)	(0.041)	(0.035)	(0.035)
Observations	2895	1861	2895	1861	2895	1861
R-sq.	0.066	0.053	0.052	0.076	0.049	0.058
Sample	All	Childless only	All	Childless only	All	Childless only

The table shows OLS estimates of the correlation between *having an income*, *stable relationships*, *childbearing*, and whether the reposndent would be happy to get a child, whether she thinks her parents would be happy and whether she thinks people would treat her with more respect. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$. All the estimations include the covariates listed in 1.

3.3. Well-being

We have seen that the interventions led to changes in self-employment and early pregnancy, which are the key outcomes targeted by the programs. Ultimately, however, we are interested in the participants' well-being, and we address this issue by using a broader measure of income, which in addition to self-employment also includes salaried work and paid work in the family business, and by showing impact on happiness, measured by the degree to which the respondents agreed with the statement "I am very happy with my life". Did the move into self-employment, and out of salaried work, actually increase overall income? And did the empowerment programs succeed in making the young women more happy with their lives?

Table 8 shows that there is a positive impact of the economic empowerment program on long-term total income. This suggests that the occupational change induced by the economic empowerment program has paid off in income terms.

The results on happiness are, however, more complex. In the short-term, all treatments caused an increase in happiness. In the mid term, these positive treatment effects have been weakened, and in the long term, for the health program and the combined program, they have been reversed.

In order to shed light on the mechanisms behind the reversal of happiness, it is useful to consider how the level of happiness evolves over time in our sample. Figure 4 shows the proportion of respondents who agree or strongly agree with the happiness statement in the different treatment arms, by survey. We see two striking patterns. First, the happiness levels are stable and high in the short- and medium term, but they drop sharply in the long term. Second, as shown in Table 8, the empowerment programs lead to increased happiness in the short term but (for the health and combined treatment) to an even larger fall in the long term.

Research on happiness provides us with at least three candidate explanations for the pattern that we observe on happiness.

First, children are associated with lower levels of happiness (Glass et al., 2016). This

Table 8: Impacts on well-being.

	(1)	(2)	(3)	(4)	(5)	(6)
	Long term income			Happiness		
	self.-emp.	wage	total	short-term	mid-term	long-term
Health	.24	-.217	.297	$.163^{**}_{\star}$	.121	$-.109^{*}$
	(.298)	(.24)	(.296)	(.08)	(.087)	(.064)
Economic	$.796^{**}_{\star\star}$	-.235	$.82^{***}_{\star\star}$	$.178^{**}_{\star}$	-.018	.026
	(.305)	(.262)	(.305)	(.08)	(.055)	(.066)
Health & Econ.	$.83^{**}_{\star\star}$	-.315	$.978^{***}_{\star\star}$	$.343^{***}_{\star\star\star}$	$.246^{***}_{\star\star\star}$	$-.213^{***}_{\star\star\star}$
	(.323)	(.249)	(.333)	(.061)	(.054)	(.066)
Mean Control	2.359	1.582	2.542	0	0	0
Obs.	3249	3044	3252	2895	2952	3249

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ."). The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

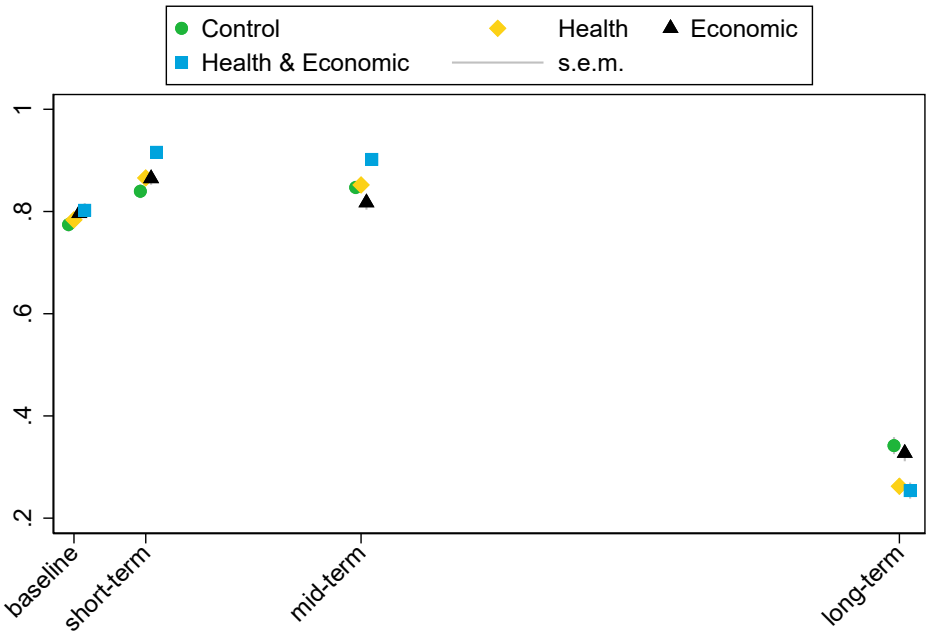
could potentially explain the negative trend in happiness over time, but it cannot fully explain the long-term negative treatment effects: the economic empowerment program also leads to more children but has no negative effect on long-term happiness.

Second, happiness levels adjust to a new reference level after a positive shock and return to normal in the longer run (Galiani et al., 2018). This theory of *hedonic adaptations* is however not a sufficient explanation since the happiness level do not only converge but are significantly *lower* among those assigned to the health intervention.

Third, the health program creates aspirations which are not fulfilled once the girls leave school, thus leading to frustration (Genicot and Ray, 2017, 2020). In particular, we know that the health program led to an increase in the girls' sense of self-control in the short term, but we also see from our data that this positive effect completely vanishes in the long term. Hence, while aspirations may be an important motivator for change (Dalton et al.; Bernard et al., 2019), there is a risk of "overshooting" such that aspirations actually become detrimental La Ferrara (2019).

This interpretation is also supported by the fact that we observe very strong negative correlations, at the individual level, between happiness in the long-term and in previous

surveys: the participants who reported high happiness levels in the short and medium terms are the ones reporting low happiness in the long-term.



Note: The figure shows the proportions of respondents who agree or strongly agree with the statement "I am very happy with my life".

Figure 4: Happiness.

4. Conclusion

We study the effect of an empowerment program targeting adolescent girls in rural Tanzania, aimed at increasing opportunities by providing new knowledge and strengthening decision making power by building self-confidence and challenging gender norms. The program consisted of an economic empowerment arm and a health empowerment arm. An important feature of our research is that we are able to distinguish the impact of the two arms and explore potential complementarities. Our project is also distinguished by its scale (involving 3500 girls), its length (tracking the participants for four years after the intervention), and its reliance on local resources (implemented at the schools, taught by local teachers).

We find that the economic empowerment intervention leads to a large and enduring positive impact on business startups and income. This is encouraging, given that the literature on business training typically finds only very muted effects from such interventions, and we suggest that it may be due to the fact that we target younger, unmarried women who are not yet restricted by family obligations.

More surprisingly, we do not find that the interventions reduce fertility. On the contrary, both interventions lead to an increase in early pregnancies. We ascribe this to the increase in income caused by economic empowerment, making it more affordable to have a child early, and to the more stable partnerships encouraged by the health program.

Our research carries an important message to policy makers about the potential, but also the complexity, of female empowerment interventions in low-income contexts. It is also a reminder to researchers to take a broad perspective when evaluating the impact of interventions, such as including measures of fertility when evaluating economic empowerment programs.

References

Almå s, I., Berge, L.I.O., Bjorvatn, K., Somville, V., Tungodden, B., 2020. Adverse selection into competition: Evidence from a large-scale field experiment in tanzania. NHH Dept. of Economics Discussion Paper 19/2020.

Ashraf, Q., Galor, O., 2011. Dynamics and Stagnation in the Malthusian Epoch. American Economic Review 101, 2003–2041.

Bailey, M.J., 2006. More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply. The Quarterly Journal of Economics 121, 289–320.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M., 2020. Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa. American Economic Journal: Applied Economics 12.

Becker, G.S., 1960. An Economic Analysis of Fertility, in: Roberts, G.B. (Ed.), Demographic and Economic Change in Developed Countries. Columbia University Press. chapter 7, pp. 209–240.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society 57, 289–300.

Berge, L.I.O., Bjorvatn, K., Maalim, A.M., Somville, V., Tungodden, B., 2018. Reducing early pregnancy in low-income countries: A literature review and new evidence, in: Anderson, S., Beaman, L., Platteau, J. (Eds.), Towards Gender Equity in Development. Oxford University Press. chapter 7.

Berge, L.I.O., Bjorvatn, K., Tungodden, B., 2014. Human and financial capital for microenterprise development: Evidence from a field and lab experiment. Management Science 61, 707–722.

Berge, L.I.O., Bjorvatn, K., Tungodden, B., 2015. Human and financial capital for microenterprise development: Evidence from a field and lab experiment. Management Science 61, 707–722.

Bernard, T., Dercon, S., Orkin, K., Taffesse, A.S., 2019. Parental aspirations for children's education: Is there a "girl effect"? experimental evidence from rural ethiopia. AEA Papers and Proceedings 109, 127–32.

Buchmann, Ninaand Field, E., Glennerster, R., Nazneen, S., Pimkina, S., Sen, I., 2018. Power vs money: Alternative approaches to reducing child marriage in bangladesh, a randomized control tria. mimeo .

Buehren, N., Goldstein, M., Gulesci, S., Sulaiman, M., Yam, V., 2017. Evaluation of an adolescent development program for girls in tanzania. World Bank Policy Research Working Papers 7961.

Dalton, P.S., Ghosal, S., Mani, A., . Poverty and Aspirations Failure. Economic Journal 126, 165–188.

Dhar, D., Jain, T., Jayachandran, S., 2020. Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India.

Doepke, M., Kindermann, F., 2019. Bargaining over babies: Theory, evidence, and policy implications. American Economic Review 109, 3264–3306.

Doepke, M., Tertilt, M., 2018. Women's empowerment, the gender gap in desired fertility, and fertility outcomes in developing countries. AEA Papers and Proceedings 108, 358–62.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association 9, 522–550.

Duflo, E., 2012. Women empowerment and economic development. Journal of Economic Literature 50, 1051–79.

Duflo, E., Dupas, P., Kremer, M., 2015. Education, HIV, and Early Fertility: Experimental Evidence from Kenya. The American Economic Review 105, 2757–2797.

Dupas, P., 2011. Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. American Economic Journal: Applied Economics 3, 1–34.

Dupas, P., Huillery, E., Seban, J., 2018. Risk information, risk salience, and adolescent sexual behavior: Experimental evidence from cameroon. Journal of Economic Behavior Organization 145, 151–175.

Field, E., Ambrus, A., 2008. Early marriage, age of menarche, and female schooling attainment in bangladesh. Journal of Political Economy 116, 881–930.

Field, E., Jayachandran, S., Pande, R., 2010. Do traditional institutions constrain female entrepreneurship? a field experiment on business training in india. American Economic Review 100, 125–29.

Galiani, S., Gertler, P.J., Undurraga, R., 2018. The Half-Life of Happiness: Hedonic Adaptation in the Subjective Well-Being of Poor Slum Dwellers to the Satisfaction of Basic Housing Needs. Journal of the European Economic Association 16, 1189–1233.

Galor, O., Weil, D.N., 2000. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. American Economic Review 90, 806–828.

Genicot, G., Ray, D., 2017. Aspirations and inequality. Econometrica 85, 489–519.

Genicot, G., Ray, D., 2020. Aspirations and economic behavior. Annual Review of Economics 12, 715–746.

Glass, J., Simon, R.W., Andersson, M.A., 2016. Parenthood and happiness: Effects of work-family reconciliation policies in 22 oecd countries. American Journal of Sociology 122, 886–929.

Goldin, C., Katz, L.F., 2002. The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. Journal of Political Economy 110, 730–770.

Hayes, R.J., Moulton, L.H., 2012. Cluster randomised trials. CRC Press, Boca Raton.

Heath, R., Mobarak, M.A., 2015. Manufacturing growth and the lives of Bangladeshi women. Journal of Development Economics 115, 1–15.

Heckman, J.J., Mosso, S., 2014. The Economics of Human Development and Social Mobility. Annual Review of Economics 6, 689–733.

Herrera, C., Sahn, D.E., Villa, K.M., 2019. Teen Fertility and Female Employment Outcomes: Evidence from Madagascar. Journal of African Economies 28, 277–303.

Hotz, J.V., Klerman, J.A., Willis, R.J., 1997. Chapter 7 The economics of fertility in developed countries. Handbook of Population and Family Economics 1, 275–347.

Jensen, R., 2012. Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India. The Quarterly Journal of Economics 127, 753–792.

Jones, L.E., Schoonbroodt, A., Tertilt, M., 2010. Fertility Theories: Can they explain the Negative Fertility-Income Relationship?, in: Shoven, J.B. (Ed.), Demography and the Economy. University of Chicago Press. chapter 2, pp. 43–100.

Karlan, D., Valdivia, M., 2011. Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. Review of Economics and Statistics 93, 510–527.

La Ferrara, E., 2019. Presidential Address: Aspirations, Social Norms, and Development. Journal of the European Economic Association 17, 1687–1722.

Lagerlöf, N.P., 2015. Malthus in Sweden. The Scandinavian Journal of Economics 117, 1091–1133.

Lundborg, P., Plug, E., Rasmussen, A.W., 2017. Can Women Have Children and a Career? IV Evidence from IVF Treatments. American Economic Review 107, 1611–37.

McKenzie, D., Woodruff, C., 2014. What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World? The World Bank Research Observer 29, 48–82.

Miller, G., 2010. Contraception as Development? New Evidence from Family Planning in Colombia. Economic Journal 120 (545), 709–736.

Mookherjee, D., Prina, S., Ray, D., 2012. A Theory of Occupational Choice with Endogenous Fertility. American Economic Journal: Microeconomics 4, 1–34.

Pearlin, L.I., Schooler, C., 1978. The structure of coping. Journal of health and social behavior , 2–21.

Rasul, I., 2008. Household bargaining over fertility: Theory and evidence from Malaysia. Journal of Development Economics 86, 215–241.

Schultz, P.T., 1997. Chapter 8 Demand for children in low income countries. Handbook of Population and Family Economics 1, 349–430.

A. Attrition

We report in Table 9 the ordinary least square estimates of the treatments effect on the probability to be included in the different surveys. We find that the attrition is not significantly correlated with the treatment assignments.

Table 9: Attrition by treatment arm in each survey.

	(1)	(2)	(3)	(4)
	Short-term	Mid-term	Long-term	
			Phone	Face-to-face
Health	-0.051	0.002	0.012	-0.000
	(0.033)	(0.027)	(0.016)	(0.028)
Economic	-0.012	-0.025	-0.004	-0.022
	(0.038)	(0.031)	(0.022)	(0.036)
Health & Economic	-0.027	-0.013	-0.000	-0.012
	(0.037)	(0.025)	(0.017)	(0.029)
Mean control	0.85	0.87	0.93	0.84
Observations	3483	3483	3483	3483

The table shows OLS estimates of the treatment impacts ("Health", "Economic" and "Health & Econ.") on the probability to be surveyed at different points in time. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. None of the coefficients are statistically significantly different from zero at the level of 0.05.

B. Heterogeneity analysis

In this section, we test whether the treatments had differential impacts along the four dimensions that we pre-specified. We selected dimensions at the individual level —age and cognition— that could in principle directly influence the students' understanding of the program, fertility and economic choices. We also selected one dimension at the household level —wealth— and one at the school level —geographical remoteness—. Our thinking was that richer household may have different opportunities to set-up a business (e.g. better access to capital) or to delay marriages and pregnancies, and that more remote schools are located in environments that are less conducive to income generation and women's independence more generally.

We estimate:

$$Y_{ij} = \alpha + \beta_1 E_j + \beta_2 H_j + \beta_3 EH_j + \delta_1 E_j * Z_{ij} + \delta_2 H_j * Z_{ij} + \delta_3 EH_j * Z_{ij} + \theta Z_{ij} + \gamma X_{ij} + \epsilon_{ij} \quad (2)$$

Where Z_{ij} is one of the following baseline variables: (i) whether school j is remote or not, (ii) whether i's household is in the top half of the wealth index or not, (iii) whether i is in the top half of the cognition index or not and (iv) whether i is more or less than 17 years old.

In the following Tables we present the estimates and standard errors of δ_1, δ_2 and δ_3.

The differential impacts on the main outcomes (self-employment, business income, teenage pregnancy and childbearing) are shown in Tables 10 and 11. The variables used to discuss the mechanisms are in Table 12 (opportunities and decision-making power), 13 (economic plans and occupations) and 14 (sexual behavior). The differential effects on long-term evolution of empowerment and happiness are reported in Table 15.

When adjusting the p-values for multiple hypothesis testing, we don't find any significant heterogeneous impacts in any of those tables except for one: the economic program alone increased the likelihood that students who were older at baseline (above 17 years of age) are still students at endline. Overall we conclude that the treatments have very homogeneous effects along the dimensions that we considered.

Table 10: Heterogeneous impacts on business development.

	(1)	(2)	(3)	(4)	(5)	(6)
	Self-employment			Business income (ihst)		
	short-term	mid-term	long-term	short-term	mid-term	long-term
Health						
Remote school	.035**	.036	-.03	.387**	.647*	.212
	(.016)	(.052)	(.064)	(.161)	(.377)	(.599)
Wealthy hh.	.009	-.016	.005	.049	-.591*	.527
	(.022)	(.037)	(.056)	(.214)	(.334)	(.48)
High cognition	-.004	-.001	.02	-.018	.281	.372
	(.017)	(.037)	(.05)	(.168)	(.309)	(.462)
Age > 17	.01	-.002	.023	.039	.149	.301
	(.024)	(.045)	(.052)	(.225)	(.342)	(.504)
Economic						
Remote school	-.009	.058	-.024	-.037	.878*	.268
	(.024)	(.063)	(.072)	(.202)	(.474)	(.604)
Wealthy hh.	.017	.02	.018	.12	-.403	.258
	(.025)	(.045)	(.054)	(.244)	(.401)	(.487)
High cognition	.02	.015	.05	.194	.21	.774*
	(.021)	(.043)	(.048)	(.193)	(.35)	(.459)
Age > 17	-.016	-.005	-.055	-.207	.261	-.182
	(.024)	(.04)	(.058)	(.22)	(.38)	(.555)
Health & Econ.						
Remote school	.015	-.013	-.015	.178	.227	.415
	(.033)	(.08)	(.068)	(.299)	(.64)	(.666)
Wealthy hh.	.043	.03	.044	.385	.267	.963**
	(.032)	(.056)	(.045)	(.292)	(.558)	(.404)
High cognition	-.012	.073	.018	-.059	.734*	.174
	(.028)	(.045)	(.051)	(.262)	(.42)	(.451)
Age > 17	.013	-.023	.042	.128	.135	.711
	(.027)	(.042)	(.051)	(.241)	(.357)	(.519)
Mean Control	.038	.187	.294	.344	1.162	2.359
Obs.	2895	2994	3249	2895	2994	3249

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 11: Heterogeneous impacts on fertility.

	(1)	(2)	(3)	(4)
	Teenage pregnancy		Started childbearing	
	mid-term	long-term	mid-term	long-term
Health				
Remote school	-.007	.007	-.007	.052
	(.022)	(.032)	(.034)	(.052)
Wealthy hh.	-.028	-.063**	-.033	-.003
	(.019)	(.031)	(.03)	(.045)
High cognition	.047*	.066*	.008	.018
	(.024)	(.035)	(.026)	(.058)
Age > 17	.005	-.009	-.033	-.047
	(.02)	(.034)	(.021)	(.043)
Economic				
Remote school	-.003	.034	.002	.012
	(.02)	(.031)	(.031)	(.054)
Wealthy hh.	-.003	.001	-.006	.015
	(.024)	(.033)	(.027)	(.053)
High cognition	.031	-.001	.017	-.042
	(.023)	(.032)	(.024)	(.057)
Age > 17	-.006	-.072**	-.031	-.1***
	(.022)	(.035)	(.022)	(.035)
Health & Econ.				
Remote school	-.021	.029	.021	-.003
	(.022)	(.034)	(.038)	(.061)
Wealthy hh.	-.015	-.058*	-.025	-.046
	(.022)	(.032)	(.034)	(.049)
High cognition	.008	.003	.013	-.027
	(.023)	(.036)	(.027)	(.06)
Age > 17	.009	-.001	.002	.024
	(.019)	(.034)	(.027)	(.042)
Mean Control	.032	.085	.056	.326
Obs.	3249	3249	2993	3262

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$ when the p-values are not corrected and by $\star \; p < 0.1$, $\star\star \; p < 0.05$, $\star\star\star \; p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 12: Heterogeneous impacts on opportunities & decision-making power.

	Panel A: Opportunities		**Panel B: Decision-Making Power**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Health	Economic	Self-control	Gender equality		Business mindset	
			index	Economic	Health	Risk	Competitiveness
Health							
Remote school	-.05	.018	.11	-.242	.139	-.508	.068
	(.032)	(.023)	(.091)	(.146)	(.188)	(.498)	(.113)
Wealthy hh.	.01	-.041*	.024	-.044	-.053	-.353	-.03
	(.025)	(.023)	(.073)	(.119)	(.135)	(.387)	(.076)
High cognition	0	-.012	.045	.054	-.073	.193	.076*
	(.017)	(.019)	(.054)	(.125)	(.133)	(.37)	(.042)
Age > 17	-.002	.033	.011	.026	.327$^{***}_{\star}$	.492	-.026
	(.018)	(.021)	(.054)	(.112)	(.111)	(.365)	(.052)
Economic							
Remote school	-.012	.013	.131	.068	.191	-.272	.065
	(.028)	(.02)	(.088)	(.157)	(.139)	(.437)	(.09)
Wealthy hh.	.026	-.029	.036	-.001	-.057	-.432	-.015
	(.025)	(.022)	(.071)	(.127)	(.111)	(.348)	(.061)
High cognition	-.017	-.019	-.022	-.206**	.088	.239	.068
	(.018)	(.021)	(.056)	(.103)	(.146)	(.327)	(.05)
Age > 17	.014	.02	.03	.159	.375$^{***}_{\star}$	.512	-.031
	(.024)	(.018)	(.054)	(.124)	(.123)	(.34)	(.06)
Health & Econ.							
Remote school	-.04	.004	.174*	-.074	.297	-.208	.166**
	(.032)	(.027)	(.103)	(.158)	(.185)	(.438)	(.083)
Wealthy hh.	.042*	-.017	.014	-.056	-.131	-.081	-.042
	(.021)	(.021)	(.07)	(.113)	(.142)	(.299)	(.067)
High cognition	-.013	.015	.131**	-.077	.047	.536*	.038
	(.018)	(.022)	(.06)	(.098)	(.15)	(.295)	(.051)
Age > 17	-.011	.001	.035	.075	.255*	.156	-.037
	(.025)	(.019)	(.065)	(.11)	(.132)	(.336)	(.056)
Mean Control	.684	.38	-2.477	3.817	3.578	7.755	.332
Obs.	2898	2898	2891	2895	2890	2895	2912

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 13: Heterogeneous impacts on economic plans & occupations

	Short-term: Plans		Long-term: Occupations				
	(1) Business	(2) Education	(3) Self-employed	(4) Wage	(5) Family business	(6) Student	(7) Domestic chores
Health							
Remote school	.088*	.034	-.03	-.033	-.011	-.019	.094*
	(.05)	(.108)	(.064)	(.032)	(.033)	(.05)	(.053)
Wealthy hh.	.012	-.001	.005	-.022	-.018	.023	.013
	(.051)	(.073)	(.056)	(.03)	(.029)	(.041)	(.042)
High cognition	.019	.037	.02	.005	.02	.021	-.067
	(.038)	(.039)	(.05)	(.031)	(.027)	(.044)	(.053)
Age > 17	.006	.033	.023	-.029	-.07**	.068*	.009
	(.042)	(.051)	(.052)	(.029)	(.033)	(.035)	(.051)
Economic							
Remote school	.005	-.09	-.024	-.055	.016	.018	.046
	(.07)	(.103)	(.072)	(.033)	(.029)	(.056)	(.053)
Wealthy hh.	.1**	.145**	.018	.013	-.016	.007	-.021
	(.05)	(.07)	(.054)	(.03)	(.029)	(.035)	(.045)
High cognition	.013	.017	.05	-.009	-.005	-.018	-.018
	(.043)	(.046)	(.048)	(.03)	(.025)	(.039)	(.055)
Age > 17	.014	.047	-.055	-.02	-.007	.14***	-.057
	(.05)	(.048)	(.058)	(.026)	(.03)	(.029)	(.047)
Health & Econ.							
Remote school	.185**	.008	-.015	-.05	-.029	.04	.054
	(.079)	(.099)	(.068)	(.032)	(.033)	(.063)	(.053)
Wealthy hh.	-.016	.02	.044	-.018	-.028	-.024	.026
	(.067)	(.064)	(.045)	(.031)	(.032)	(.039)	(.047)
High cognition	.051	.027	.018	-.012	.025	.022	-.053
	(.054)	(.061)	(.051)	(.032)	(.03)	(.046)	(.052)
Age > 17	-.023	.102*	.042	-.011	-.053*	.021	0
	(.051)	(.052)	(.051)	(.025)	(.028)	(.039)	(.042)
Mean Control	.151	.694	.294	.11	.07	.191	.335
Obs.	2894	2892	3249	3249	3249	3249	3249

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$ when the p-values are not corrected and by $\star$ $p < 0.1$, $\star\star$ $p < 0.05$, $\star\star\star$ $p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 14: Heterogeneous impacts on sexual behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Short-term: Plans		Stable relationship		Number of	Condom	Unwanted
	marriage	1st child	mid-term	long-term	partners	use	sex
Health							
Remote school	-.493	-.414	.023	.032	.409*	.032	.004
	(.436)	(.446)	(.076)	(.042)	(.246)	(.053)	(.031)
Wealthy hh.	.18	.296	-.109*	-.069	-.363**	.025	.007
	(.367)	(.331)	(.061)	(.045)	(.176)	(.048)	(.029)
High cognition	-.392	-.727**	-.002	.015	-.057	-.035	.055*
	(.36)	(.307)	(.049)	(.041)	(.157)	(.037)	(.031)
Age > 17	.257	.34	-.005	-.055	-.254	-.038	-.019
	(.3)	(.357)	(.057)	(.044)	(.168)	(.038)	(.03)
Economic							
Remote school	-.748	-.51	-.048	.087**	.153	.037	-.015
	(.459)	(.444)	(.058)	(.041)	(.208)	(.058)	(.033)
Wealthy hh.	.396	.504	-.104**	-.085*	-.2	.013	.024
	(.339)	(.358)	(.042)	(.046)	(.153)	(.056)	(.026)
High cognition	-.506	-.622	.046	-.046	-.027	-.004	.027
	(.392)	(.396)	(.053)	(.04)	(.137)	(.046)	(.023)
Age > 17	.336	.558*	-.047	-.067	-.404**	.036	-.003
	(.248)	(.319)	(.04)	(.048)	(.168)	(.046)	(.027)
Health & Econ.							
Remote school	-.947**	-.8*	-.035	-.031	-.166	-.024	-.024
	(.436)	(.408)	(.072)	(.042)	(.218)	(.071)	(.032)
Wealthy hh.	.485	.365	-.038	-.03	-.064	.002	.038
	(.292)	(.247)	(.049)	(.046)	(.167)	(.057)	(.029)
High cognition	.199	-.27	.004	-.04	-.072	-.029	.016
	(.362)	(.377)	(.049)	(.047)	(.15)	(.041)	(.03)
Age > 17	.487*	.581	.005	-.044	-.405**	.012	-.002
	(.283)	(.357)	(.055)	(.047)	(.168)	(.051)	(.033)
Mean Control	25.584	26.737	.286	.754	1.8	.815	.101
Obs.	2895	2863	2722	2895	2895	2506	2506

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

Table 15: Heterogeneous impacts on well-being.

	(1)	(2)	(3)	(4)	(5)	(6)
	Happiness			Long term income		
	short-term	mid-term	long-term	self.-emp.	wage	total
Health						
Remote school	.212	-.41	.391	.039	-.089	-.147
	(.599)	(.436)	(.587)	(.16)	(.176)	(.131)
Wealthy hh.	.527	-.391	.394	.142	.092	.069
	(.48)	(.421)	(.509)	(.12)	(.152)	(.133)
High cognition	.372	.27	.433	.13	-.119	-.148
	(.462)	(.396)	(.456)	(.095)	(.105)	(.11)
Age > 17	.301	-.14	.302	-.078	.092	.14
	(.504)	(.356)	(.498)	(.108)	(.116)	(.123)
Economic						
Remote school	.268	-.851*	.24	.092	.119	-.047
	(.604)	(.436)	(.607)	(.155)	(.102)	(.138)
Wealthy hh.	.258	.104	.22	-.031	.157	-.131
	(.487)	(.357)	(.518)	(.113)	(.103)	(.12)
High cognition	.774*	-.002	.784*	.156	0	-.179*
	(.459)	(.381)	(.438)	(.124)	(.1)	(.09)
Age > 17	-.182	-.211	-.335	.044	.058	.32***
	(.555)	(.286)	(.558)	(.118)	(.13)	(.12)
Health & Econ.						
Remote school	.415	-.638	.436	.139	.126	-.024
	(.666)	(.466)	(.7)	(.118)	(.105)	(.134)
Wealthy hh.	.963**	-.164	.72	-.025	-.071	.039
	(.404)	(.371)	(.461)	(.11)	(.109)	(.119)
High cognition	.174	-.237	0	.156*	-.02	-.142
	(.451)	(.465)	(.471)	(.09)	(.115)	(.093)
Age > 17	.711	-.087	.469	.056	.064	.14
	(.519)	(.382)	(.512)	(.099)	(.125)	(.113)
Mean Control	2.359	1.582	2.542	0	0	0
Obs.	3249	3044	3252	2895	2952	3249

The table shows OLS estimates of the interaction between the treatment assignment and the baseline variable of interest. The standard errors, in parenthesis, are clustered at the level of randomization, the schools. Statistically significant differences between the estimates and zero are indicated by $* \ p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$ when the p-values are not corrected and by $\star \ p < 0.1$, $\star\star \ p < 0.05$, $\star\star\star \ p < 0.01$ when the p-values are corrected for multiple hypothesis testing. All the estimations include the covariates listed in 1.

C. Multiple hypothesis testing

In Tables 2, 3, 4, 5, 8, 10, 11, 12, 13, 14 and 15, we provide the results of hypotheses' tests based on both, unadjusted p-values and p-values adjusted for multiple hypothesis testing. To adjust the p-values, we control the false discovery rate by following the procedure described in Benjamini and Hochberg (1995).

We adjust the p-values for the fact that we perform multiple tests per outcome variable. We first estimate the equation:

$$Y_{ij} = \alpha + \beta_1 E_j + \beta_2 H_j + \beta_3 EH_j + \gamma X_{ij} + \epsilon_{ij} \quad (1)$$

We then ask three questions:

1. Compared to the control group, does each treatment has an impact on outcome Y_{ij}? That is, is β_1 equal to zero? Is β_2 equal to zero? And is β_3 equal to zero?

2. Do the three treatment arms have impacts that are equal to each other? That is, is β_1 equal to β_2? Is β_1 equal to β_3? And is β_2 equal to β_3?

3. Is the impact of the combined treatement different from the sum of the separate treatments? That is, is $\beta_1 + \beta_2$ equal to β_3?

And we adjust the p-values *per question*.

We also adjust the p-values for the fact that we test the treatment impacts on several related outcomes. We group the outcomes into XX families and correct the p-values within families. The self-employment and business income variables of Table 2 make one family. The variables of Table 3 are grouped in another family. In table 4, we group in one family the opportunities variables, the business mindset variables in another family and the remaining variables in yet another family. The short-term plans of Table 5 are in one family and the long-term occupations in another one, the varibales of panel B are all grouped together. Finally, in Table 8 we group together the happiness measures in one family and the income measures in another.

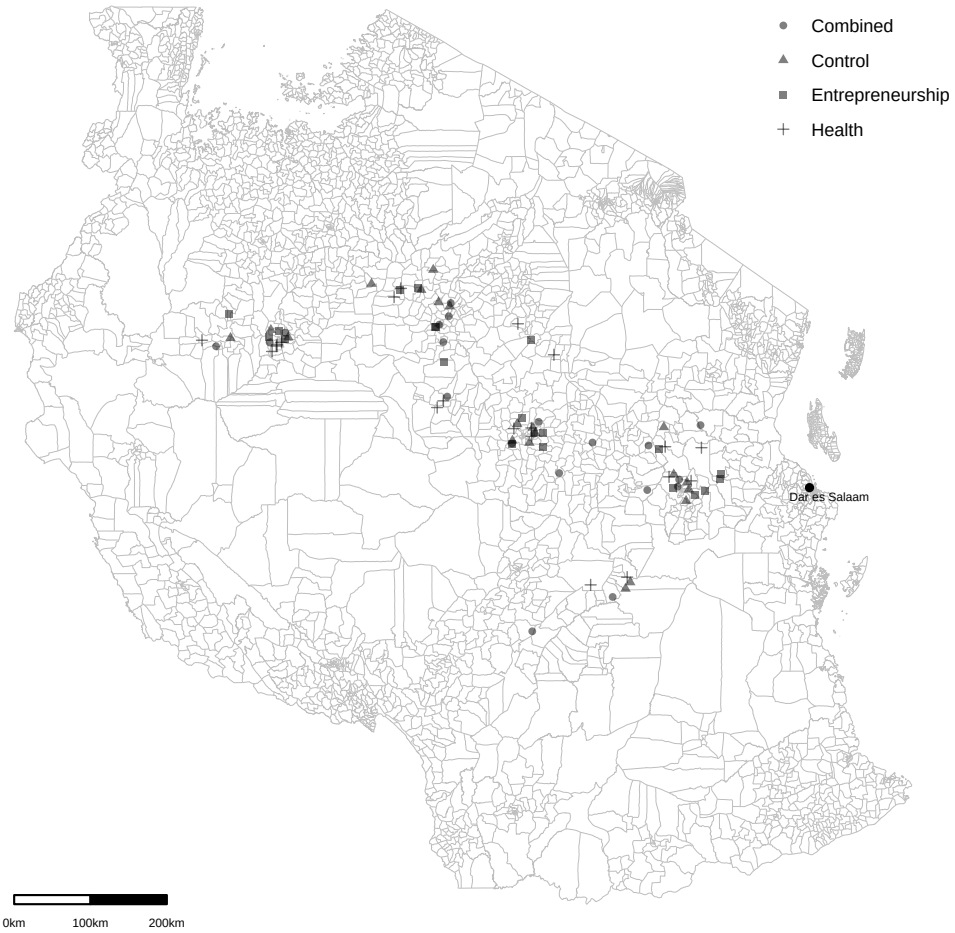
To illustrate, in Table 2 we first group together the 18 p-values corresponding to the columns (1) to (6) (first group of outcomes) and the rows "Health", "Economic" and "Health & Econ." (our first question), and we apply the procedure of Benjamini and Hochberg (1995) to these 18 p-values. We then group together the six p-values corresponding to the columns (1) to (6) and the row "Health + Econ. - Health & Econ." (third question) and adjust those six p-values. Then we group together the 18 p-values corresponding to the columns (1) to (6) and the rows "Econ. - Health", "Econ. - Health & Econ." and "Health - Health & Econ." (second question) and adjust those 18 p-values.

D. Variables

Do make an appendix with the definitions of all the variables?

E. Map of the study sites

Figure 5: Map of the study sites.



The map shows the borders of all the Tanzanian wards and the locations of the schools that participated in the study.

F. Curriculum

Each of the programs were made of eight modules. The economic empowerment program, "*Build your life*", included:

1. We are girls, we can!
2. Being an entrepreneur.
3. Business ideas and different types of businesses to start.
4. Marketing and customer care.
5. Resources you will need.
6. Business, security and relationships.
7. How to think about money.
8. Planning your business and moving forward

and the health empowerment program, "*Protect your life*", included:

1. We are girls.
2. Coming of age.
3. Healthy relationships.
4. Let's talk about sex.
5. Staying safe: part I.
6. Staying safe: part II.

7. Violence against women.

8. Moving forward.

G. Pre-analysis plans

refer to PAPs and clarify un-planned analysis Do we also want to add all the het effects tables?

Table 16: Pre-analysis plan: short-term impacts.

Knowledge

	Health knowledge		Business knowledge	
	(1)	(2)	(3)	(4)
Health	0.058	0.027	-0.051	0.002
	(0.120)	(0.125)	(0.061)	(0.059)
Entrepreneurship	0.020	0.022	0.126**	0.164***
	(0.110)	(0.120)	(0.061)	(0.056)
Health &	0.143	0.115	0.267***	0.308***
Entrepreneurship	(0.126)	(0.121)	(0.065)	(0.071)
Mean control	4.760	4.760	1.910	1.910
Observations	3345	2898	3346	2898
Controls	No	Yes	No	Yes

Behavior

	Safe sex		Business plans	
	(5)	(6)	(7)	(8)
Health	-0.058	-0.042	0.012	0.022
	(0.035)	(0.034)	(0.030)	(0.028)
Entrepreneurship	-0.079	-0.057	0.364***	0.378***
	(0.066)	(0.045)	(0.040)	(0.039)
Health &	-0.025	-0.013	0.403***	0.432***
Entrepreneurship	(0.036)	(0.037)	(0.045)	(0.046)
Mean control	0.685	0.685	0.161	0.161
Observations	3347	2894	3348	2894
Controls	No	Yes	No	Yes

Table 16: (continued)

Table 16: (continued)

Gender equality

	Acceptance of wife beating		Wife earner	
	(9)	(10)	(11)	(12)
Health	-0.049**	-0.056***	0.156*	0.136*
	(0.020)	(0.019)	(0.078)	(0.077)
Entrepreneurship	0.018	0.018	0.282***	0.297***
	(0.017)	(0.014)	(0.075)	(0.079)
Health &	-0.053***	-0.063***	0.253***	0.302***
Entrepreneurship	(0.019)	(0.018)	(0.077)	(0.076)
Mean control	0.283	0.283	3.820	3.820
Observations	3342	2889	3349	2895
Controls	No	Yes	No	Yes

Empowerment

	Compete		Empowerment (-)	
	(13)	(14)	(15)	(16)
Health	-0.025	-0.027	-0.129**	-0.126***
	(0.057)	(0.057)	(0.055)	(0.047)
Entrepreneurship	0.051	0.063	-0.071	-0.057
	(0.047)	(0.048)	(0.051)	(0.041)
Health &	0.006	0.010	-0.187***	-0.181***
Entrepreneurship	(0.047)	(0.044)	(0.056)	(0.051)
Mean control	0.321	0.321	2.499	2.499
Observations	3362	2912	3343	2891
Controls	No	Yes	No	Yes

Table 16: (continued)

The Table displays the estimated OLS coefficients and their standard errors in parenthesis. The standard errors are clustered at the level of randomization (the schools).

Table 17: Pre-analysis plan: Mid-term impacts.

Behavior

	Childbearing		Business Startup	
	(1)	(2)	(3)	(4)
Health	0.014	0.008	0.044	0.037
	(0.016)	(0.017)	(0.030)	(0.025)
Entrepreneurship	0.008	0.008	0.125***	0.114***
	(0.014)	(0.014)	(0.034)	(0.032)
Health &	0.018	0.021	0.183***	0.167***
Entrepreneurship	(0.016)	(0.017)	(0.039)	(0.038)
Mean control	0.056	0.056	0.185	0.185
Observations	3341	2993	3342	2994
Controls	No	Yes	No	Yes

Behavior (continued)

	Business revenues		Patience	
	(5)	(6)	(7)	(8)
Health	0.629	0.542	-0.040	-0.041
	(0.391)	(0.375)	(0.034)	(0.031)
Entrepreneurship	1.524***	1.347***	-0.058	-0.037
	(0.440)	(0.491)	(0.038)	(0.037)
Health &	2.385***	2.169***	-0.029	-0.035
Entrepreneurship	(0.641)	(0.608)	(0.040)	(0.035)
Mean control	1.554	1.554	0.453	0.453
Observations	3343	2994	3342	2993
Controls	No	Yes	No	Yes

Table 17: (continued)

Table 17: (continued)

Gender equality

	Acceptance of wife beating		Acceptance of wife's higher earnings	
	(9)	(10)	(11)	(12)
Health	-0.003	-0.005	0.102	0.099
	(0.004)	(0.004)	(0.084)	(0.087)
Entrepreneurship	-0.000	-0.002	0.199**	0.223**
	(0.004)	(0.004)	(0.087)	(0.097)
Health &	-0.000	0.000	0.084	0.105
Entrepreneurship	(0.005)	(0.005)	(0.091)	(0.090)
Mean control	0.009	0.009	3.458	3.458
Observations	3936	3479	3343	2994
Controls	No	Yes	No	Yes

Empowerment

	Not in control		Feeling useless	
	(13)	(14)	(15)	(16)
Health	-0.152*	-0.127	-0.054	-0.025
	(0.077)	(0.082)	(0.095)	(0.100)
Entrepreneurship	-0.105	-0.069	-0.115	-0.091
	(0.066)	(0.066)	(0.094)	(0.103)
Health &	-0.169**	-0.154*	-0.124	-0.105
Entrepreneurship	(0.083)	(0.085)	(0.102)	(0.106)
Mean control	3.456	3.456	2.612	2.612
Observations	3343	2994	3343	2994
Controls	No	Yes	No	Yes

Table 17: (continued)

Table 17: (continued)
Welfare

	Happy	
	(17)	(18)
Health	0.123	0.138
	(0.110)	(0.099)
Entrepreneurship	0.001	-0.020
	(0.067)	(0.062)
Health &	0.256***	0.279***
Entrepreneurship	(0.060)	(0.061)
Mean control	3.814	3.814
Observations	3300	2952
Controls	No	Yes

	Happy with health		Happy with econ.	
	(19)	(20)	(21)	(22)
Health	0.013	0.000	0.132	0.133
	(0.067)	(0.063)	(0.100)	(0.085)
Entrepreneurship	-0.005	-0.014	0.224**	0.204**
	(0.068)	(0.064)	(0.087)	(0.085)
Health &	0.075	0.092	0.229**	0.263***
Entrepreneurship	(0.060)	(0.059)	(0.097)	(0.093)
Mean control	4.519	4.519	2.539	2.539
Observations	3343	2994	3342	2994
Controls	No	Yes	No	Yes

Table 17: (continued)

	Sickness		Income	
	(23)	(24)	(25)	(26)
Health	0.033	0.041	-0.025	-0.201
	(0.034)	(0.032)	(0.471)	(0.473)
Entrepreneurship	0.001	0.015	0.105	-0.001
	(0.031)	(0.034)	(0.498)	(0.462)
Health &	-0.001	0.002	0.551	0.580
Entrepreneurship	(0.030)	(0.034)	(0.511)	(0.471)
Mean control	0.259	0.259	15.408	15.408
Observations	3342	2993	3343	2994
Controls	No	Yes	No	Yes

The Table displays the estimated OLS coefficients and their standard errors in parenthesis. The standard errors are clustered at the level of randomization (the schools).

Table 18: Pre-analysis plan: Long-term impacts.

Behavior

	Childbearing		Business Startup	
	(1)	(2)	(3)	(4)
Health	0.054*	0.051*	0.038	0.036
	(0.030)	(0.027)	(0.033)	(0.032)
Entrepreneurship	0.050*	0.056**	0.091**	0.093**
	(0.030)	(0.028)	(0.036)	(0.036)
Health &	0.023	0.024	0.120***	0.118***
Entrepreneurship	(0.032)	(0.030)	(0.035)	(0.032)
Mean control	0.323	0.323	0.294	0.294
Observations	3253	3249	3253	3249
Controls	No	Yes	No	Yes

Behavior (continued)

	Business revenues		Patience	
	(5)	(6)	(7)	(8)
Health	0.272	0.240	-0.000	-0.003
	(0.296)	(0.299)	(0.029)	(0.029)
Entrepreneurship	0.784**	0.796**	-0.005	-0.011
	(0.299)	(0.305)	(0.032)	(0.032)
Health &	0.835**	0.830**	-0.009	-0.015
Entrepreneurship	(0.334)	(0.324)	(0.031)	(0.031)
Mean control	2.359	2.359	0.643	0.643
Observations	3253	3249	3251	3247
Controls	No	Yes	No	Yes

Table 18: (continued)

Behavior (continued)

	Risky sexual behavior		Takes risks	
	(9)	(10)	(11)	(12)
Health	0.348*	0.323*	9.679	13.694
	(0.191)	(0.175)	(111.969)	(111.946)
Entrepreneurship	-0.088	-0.075	44.054	50.846
	(0.173)	(0.161)	(94.796)	(90.434)
Health &	-0.203	-0.212	31.737	29.764
Entrepreneurship	(0.160)	(0.151)	(73.382)	(75.396)
Mean control	-0.000	-0.000	2815.321	2815.321
Observations	2899	2895	2899	2895
Controls	No	Yes	No	Yes

Gender equality

	Acceptance of wife beating		Wife's decision power	
	(13)	(14)	(15)	(16)
Health	-0.012	-0.022	0.020	0.016
	(0.017)	(0.017)	(0.023)	(0.023)
Entrepreneurship	0.032	0.015	0.010	0.001
	(0.026)	(0.021)	(0.022)	(0.021)
Health &	-0.005	-0.011	-0.004	-0.007
Entrepreneurship	(0.020)	(0.018)	(0.026)	(0.025)
Mean control	0.099	0.099	0.149	0.149
Observations	3253	3249	2899	2895
Controls	No	Yes	No	Yes

Table 18: (continued)

Table 18: (continued)

Empowerment

	Not in control	
	(17)	(18)
Health	0.097	0.105
	(0.093)	(0.091)
Entrepreneurship	-0.110	-0.097
	(0.105)	(0.106)
Health &	0.113	0.114
Entrepreneurship	(0.092)	(0.090)
Mean control	3.557	3.557
Observations	3253	3249
Controls	No	Yes

Welfare

	Happy		Index bad health	
	(19)	(20)	(21)	(22)
Health	-0.102	-0.118*	0.017	0.008
	(0.070)	(0.069)	(0.029)	(0.025)
Entrepreneurship	0.028	0.028	-0.010	-0.024
	(0.081)	(0.071)	(0.026)	(0.025)
Health &	-0.228***	-0.230***	-0.014	-0.021
Entrepreneurship	(0.077)	(0.071)	(0.025)	(0.022)
Mean control	2.468	2.468	0.154	0.154
Observations	3253	3249	2740	2736
Controls	No	Yes	No	Yes

Table 18: (continued)

The Table displays the estimated OLS coefficients and their standard errors in parenthesis. The standard errors are clustered at the level of randomization (the schools).