



Article

Girl Empower – A gender transformative mentoring and cash transfer intervention to promote adolescent wellbeing: Impact findings from a cluster-randomized controlled trial in Liberia

Berk Özler^{a,*}, Kelly Hallman^b, Marie-France Guimond^c, Elizabeth A. Kelvin^d, Marian Rogers^c, Esther Karnley^c

^a The World Bank, 1818 H Street, NW, Washington, DC, 20433, United States

^b Population Council, United States

^c International Rescue Committee, United States

^d City University of New York, United States



ARTICLE INFO

Keywords:

Sexual violence
Mentoring programs
Cash transfers
Adolescent welfare

ABSTRACT

Background: We evaluated *Girl Empower* – an intervention that aimed to equip adolescent girls with the skills to make healthy, strategic life choices and to stay safe from sexual abuse using a cluster-randomized controlled trial with three arms: control, *Girl Empower* (GE), and GE+.

Methods: GE delivered a life skills curriculum to girls aged 13–14 in Liberia, facilitated by local female mentors. In the GE + variation, a cash incentive payment was offered to caregivers for girls' participation in the program. We evaluated the impact of the program on seven pre-specified domains using standardized indices: sexual violence, schooling, sexual and reproductive health (SRH), psychosocial wellbeing, gender attitudes, life skills, and protective factors.

Findings: Participation rates in the program were high in both GE and GE+, with the average participant attending 28 out of 32 sessions. At 24 months, the standardized effects of both GE and GE+, compared to control, on sexual violence, schooling, psychosocial wellbeing, and protective factors were small (β , ≤ 0.11 standard deviations [SD]) and not statistically significant at the 95% level of confidence. However, we found positive standardized effects on Gender Attitudes (GE: β , 0.206 SD, $p < 0.05$; GE+: β , 0.228 SD, $p < 0.05$), Life Skills (GE: β , 0.224 SD, $p < 0.05$; GE+: β , 0.289 SD, $p < 0.01$), and SRH (GE: β , 0.244 SD, $p < 0.01$; GE+: β , 0.372 SD, $p < 0.01$; F-test for GE = GE+: $p = 0.075$).

Interpretation: *Girl Empower* led to sustained improvements in several important domains, including SRH, but did not reduce sexual violence among the target population.

Violence against women (VAW) and children (VAC) are global epidemics that have lifelong impacts on the health and welfare of individuals, families, and communities. The epidemics are closely linked; VAW and VAC tend co-occur within households and exposure to VAC predicts female experience and male perpetration of intimate partner violence (IPV) during adulthood (Guedes, Bott, Garcia-Moreno, & Colombini, 2016). Because of their age, however, adolescents may not have access to supportive interventions, most of which are aimed at supporting either married adult females or younger children; adolescent girls may therefore not have access to violence prevention or response services (Ellsberg et al., 2017; Bruce & Hallman, 2008). Liberia, the

setting for this study, has a history of armed conflict during which women suffered greatly (Swiss et al., 1998; Dziewanski, 2012); high levels of interpersonal and sexual violence (SV) continue to occur, particularly in areas that saw high conflict events and fatalities during the civil war (Kelly, Colantuoni, Robinson, & Decker, 2018).

Programs designed to work specifically with adolescent girls as a unique subpopulation began to appear in the early 1990s in the HIV field and early 2000s more generally (Bruce & Hallman, 2008; Gibbs, Jacobson, & Kerr Wilson, 2017). Some of these (reviewed by (Bruce & Hallman, 2008; Marcus, Gupta-Archer, Darcy, & Page, 2017)) have shown favorable impacts on sexual health behaviors, HIV and other

* Corresponding author.

E-mail address: bozler@worldbank.org (B. Özler).

<https://doi.org/10.1016/j.ssmph.2019.100527>

Received 28 January 2019; Received in revised form 17 October 2019; Accepted 1 December 2019

Available online 12 December 2019

2352-8273/© 2020 The World Bank.

Published by Elsevier.

This is an open access article under the CC BY-NC-ND IGO license

(<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/>).

sexually transmitted infections (STIs), and child marriage. Few, however, have focused specifically on SV experienced by adolescents, particularly among those younger than 15 years of age, out-of-school, married, displaced or who have experienced multiple forms of violence (Gibbs et al., 2017; Yount, Krause, & Miedema, 2017).

Girl-focused interventions designed to reduce SV mainly attempt to address the accepted drivers of female disadvantage and victimization: poverty, low earning power, social isolation, and/or harmful gender norms. Although different categorizations have been employed in recent reviews (Lundgren & Amin, 2015; Marcus et al., 2017), such programs tend to include one or more of the following components: (a) a cash transfer, frequently conditioned on income status or a behavior; (b) economic skills strengthening without a cash transfer (vocational skills, financial education, savings, and/or microcredit); and (c) gender transformative content, usually delivered to girls by a mentor in a safe space, to guardians or community members, or occasionally to both girls and adults.

Several cash (or in-kind) transfer schemes have assessed impacts on adolescent sexual behaviors (Cluver et al., 2013, 2014; Handa et al., 2015; Heinrich, HJ, & Samson, 2017; Minnis et al., 2014; Rosenberg, Pettifor, Thirumurthy, Halpern, & Handa, 2014); and/or HIV, herpes simplex virus 2 (HSV-2), or other sexually transmitted infections (STIs) (Baird, Garfein, McIntosh, & Ozler, 2012; Bjorkman Nyqvist, Corno, de Walque, & Svensson, 2018; Duflo, Dupas, & Kremer, 2015; Kohler & Thornton, 2012; Pettifor et al., 2016; Walque et al., 2012, 2017). While the impacts on child marriage and teen pregnancy are promising, especially in the short-term, only two studies found reductions in HIV or STI risk (Baird et al., 2012; Bjorkman Nyqvist et al., 2018), and the one with a longer term follow-up revealed that the impacts of unconditional cash transfers on pregnancy, marriage, and HIV largely evaporated after the transfers stopped (Baird, McIntosh, & Özler, 2019). Cash transfer programs have also been demonstrated to reduce IPV reported by adult women (Buller et al., 2016, 2018; Gibbs et al., 2018; Haushofer, Ringdal, Shapiro, & Wang, 2019; Heath, Hidrobo, & Roy, 2018; Hidrobo, Peterman, & Heise, 2016; Roy, Hidrobo, Hoddinott, & Ahmed, 2019). To date, there is only one evaluated cash transfer program that assessed impacts on SV experienced by adolescent girls (Kilburn et al., 2018) and it reduced physical IPV among South African high school girls aged 13–20 by reducing their likelihood of being in a sexual relationship, but did not have an effect on forced sex by intimate partners. IPV, physical or forced sex, by non-intimate partners (such as family, teachers, neighbors) was not included in the outcome. A review of social safety net (SSN) programs on childhood violence used a broader definition of sexual violence (sexual exploitation and/or sexual abuse) and found that SSNs had a statistically significant protective impact on 40% of SV indicators measured (Peterman, Neijhoft, Cook, & Palermo, 2017).

Another stream of interventions for adolescent girls combines financial education and gender transformative content. This combination was inspired by an earlier wave of programs, which suggested that providing micro-loans alone might put adolescents at greater risk for experiencing SV, although the evidence for this hypothesis is based on studies that are either non-experimental or rely on small samples (Austrian & Muthengi, 2014; Dunbar et al., 2010). Financial education (as distinct from livelihoods training and/or micro-loans) combined with gender-transformative interventions, such as girls' clubs, have demonstrated favorable impacts on sexual and reproductive health (SRH) behaviors (Dunbar et al., 2014; Jewkes et al., 2014; Bandiera, Buehren, Burgess, & et al, forthcoming), and SV (Austrian & Muthengi, 2014; Dunbar et al., 2014; Jewkes et al., 2014; Bandiera et al., forthcoming), but most studies focus on older adolescents up to age 24, most of whom are already in sexual relationships.

This study examines the potential of adding a cash transfer component to a gender transformative mentoring intervention, which aimed to reduce sexual abuse among females in early adolescence. Although this approach is unlikely to shift entrenched norms and practices around sexual violence, providing girls with additional knowledge and skills via

a safe spaces' platform could increase their supportive social connections (local adult female mentors and same-age female friends), and provide protection within myriad unsafe situations of their environment. Monthly parenting sessions for caregivers were intended to bolster adult supportive attitudes toward their girls. The cash incentive for girls' attendance was primarily intended to boost program impacts by increasing participation, but also hypothesized to enable girls avoid sexual violence by delaying entry into unsafe sexual relationships (Temin, Amin, Ngo, & Psaki, 2018). The program was implemented by the International Rescue Committee (IRC) in rural Liberia and was aimed at girls, aged 13–14. The target age range was chosen based on the high rates of sexual debut (57%), school dropout (43%), pregnancy (10%), and marriage (10%) already experienced by females aged 15–17 in the study region, as reported by the Demographic and Health Surveys in 2007 (Liberia Institute of Statistics and Geo-Information Services (LISGIS) [Liberia], 2007).

Methods

The Girl Empower intervention in Liberia

Girl Empower was implemented by the IRC in 56 communities in Nimba County, Liberia from February to November 2016. It aimed to equip adolescent girls with the skills and experiences necessary to make healthy, strategic life choices and to stay safe from sexual abuse. Compared with the national average, the rate of ever experiencing sexual violence in the North Central region (where Nimba County is situated) was somewhat but not markedly higher (20.2 vs. 17.6% of adult women); young women aged 20–24 married 1.5 years younger (18.2 vs. 19.7 years); and primary school net attendance ratio for girls was lower (28.2 vs. 38.6) (Liberia Institute of Statistics and Geo-Information Services (LISGIS) [Liberia], 2007).

Girl Empower was implemented in two treatment variations, called "GE" and "GE+". Both GE and GE + consisted of 1) *Girl Empower* life skills curriculum, facilitated by local female mentors; 2) Caregiver discussion groups, facilitated by IRC staff; 3) Individual savings start-up for the girls; and 4) Capacity building for local health and psychosocial service providers. In the GE + variation, the IRC added an incentive payment to caregivers tied to girls' participation in weekly sessions.

Life skills curriculum (GE and GE+)

Girl Empower centered on a mentorship program where 65 groups of five to 20 adolescent girls, aged 13–14, met weekly with local trained female mentors, aged 20 to 35, for a total of 39 weeks. The meetings took place in safe spaces located in, and designated by, the communities. Two mentors were assigned to each group (130 mentors in total) to facilitate 32 weekly sessions based on a life skills curriculum, which covered: Sense of self; Feelings and emotions; Social networks; Protection and safety; Financial literacy; Reproductive Health; Leadership and Empowerment; and Setting life goals. The groups met for an additional seven weeks to prepare a community action event as well as the graduation ceremonies, which brought girls from intervention communities to three central locations. During the graduation ceremonies, the *Girl Empower* participants demonstrated the knowledge they acquired during the program and advocated for support for their safety and protection in their homes, schools, and community at large. The mentors and adolescent girls invited key stakeholders including community leaders, government officials and international partners to these events.

Caregiver sessions (GE and GE+)

IRC staff facilitated eight monthly sessions with 759 parents and caregivers of the program participants. These sessions aimed at familiarizing the parents and caregivers with the curriculum content, support them in reinforcing the skills that the girls learned, and encourage them to support and protect girls in their communities. The IRC provided information on the challenges faced by adolescent girls, led discussions

and reflections on how girls and boys are raised by their caregivers, including ways in which these practices may directly or indirectly cause harm, as well as on positive and supportive parenting strategies when interacting with their daughters.

Individual savings start-up (GE and GE+)

All participating girls received cash to help start their own savings account, along with a savings book and a cash box. Each girl received \$2 per month for a total of \$16 during the eight-month implementation period. IRC Liberia contracted a local NGO to implement the cash transfer scheme, which announced the day the monthly transfers would be made in each treatment community ahead of time. The savings payments were provided directly to the participating girls as cash in envelopes. An IRC assessment found that program areas did not have ready access to formal banking institutions or good network coverage for mobile transfers. As the savings payments were not tied to attendance, girls who were absent on payment day accumulated their payments and received them on the next payment day they attended. The envelopes for the girls were not given to their caregivers.

Participation incentive payment (GE + only)

In the GE + variation, caregivers of program participants received of a payment of \$1.25 for each of the 32 regular sessions that the adolescent girl attended, for a maximum of \$40. Once the previous month's attendance was known and confirmed for each participant, IRC Liberia staff relayed this information to the same local NGO in charge of the cash transfer scheme, which prepared envelopes containing the right amount of cash to be given to the caregivers. To avoid multiple trips to the same treatment community, the *participation incentive* payments to the *caregivers* were made at the same meeting as the *savings* payments to the *girls*. Absent caregivers could similarly accumulate their payments and receive them at the next payment meeting they attended. The envelopes for the caregivers were not given to the girls. The implied transfer amount of \$40 for a girl with perfect *Girl Empower* attendance was determined by IRC Liberia as an estimate of the (financial and opportunity) cost of attending school for one year, an impact they hoped *Girl Empower* would have. This amount of approximately \$6/month constitutes more than 10% of per capita consumption in Liberia (World Bank, 2018), and likely more in the study area, because Nimba is a poorer than average rural county in Liberia.

The IRC also trained 56 local health and psychosocial service providers in the treatment communities to provide quality services to survivors of gender-based violence.

Study procedures

We conducted a parallel cluster-randomized controlled trial with three arms: control, GE, and GE+ (allocation ratio: 1:1:1). In order to reach the estimated required sample size, we used data from the 2008 census of the Liberia Institute of Statistics and Geo-Information Services to develop a list of 100 villages likely to have at least five eligible girls aged 13 or 14. 10,930 households were listed in these 100 villages between July and September 2015 and 16 villages were found to contain fewer than five eligible girls and removed from the study sample. In the remaining 84 villages, we surveyed 1,216 eligible girls and 1,132 caregivers. Ex ante power calculations to detect a 10-percentage point (pp) reduction in sexual debut and a 5-pp reduction in child marriages (from expected control means of 57% and 10%, respectively, using the Liberia Demographic and Health Survey conducted in 2013⁴²(Liberia Institute of Statistics and Geo-Information Services [LISGIS] Liberia, 2014)) were used to determine sample size. The data source used for power calculations did not include data on sexual violence, which is why the power calculations were based on these secondary outcomes.

Stratified randomization at the village level was conducted to assign villages to the three study arms, where six strata were formed using the number of eligible girls and the mean level of school enrollment in each

village. Within each stratum, villages were assigned randomly to control, GE, or GE + conditions. A dofile was written in Stata 13.1 to conduct the random assignment, which was overseen by the lead author. Fig. 1 (Study Flow) describes the allocation of villages to the study arms.

The IRC team conducted registration for *Girl Empower* in the 56 treatment villages in late 2015, during which 375 (396) of the 402 (415) study participants in GE (GE+) communities were found, all of whom consented to participate in the program. In nine of the 56 treatment villages, where the number of girls registered exceeded 20 (the maximum number of girls in a group deemed manageable for the mentors by the IRC staff), two groups were created, for a total of 65 groups in 56 villages. The IRC also identified, vetted, and trained 130 female mentors. Two female mentors led each group. Only registered participants were allowed to attend *Girl Empower* sessions.

Follow-up data collection started in August 2017 and continued through February 2018 with an extensive tracking effort to minimize loss-to-follow-up, which was low (Fig. 1). We interviewed 96.7 percent of the study participants at follow-up. Similarly, of the 1,136 caregivers interviewed at baseline, 1,082 (or 95.9 percent) were successfully interviewed at follow-up.

Institutional Review Boards at the Population Council and the University of Liberia approved the study procedures. Informed assent/consent was obtained from all participants prior to enrollment in the study. The trial is registered at the American Economic Association's AEA RCT Registry (AEARCTR-0002717) and at ClinicalTrials.gov (NCT03756467).

Measures

We collected data in seven domains at baseline and the 24-month follow-up: SV (primary outcome), schooling, SRH, psychosocial well-being, gender attitudes, life skills, and protective factors (secondary outcomes). Within each of these domains, trained female enumerators fluent in all local languages administered multiple questions or scales to the study participants using tablets. The questionnaires used to collect the data are under [Supporting Documents and Materials](#) in the AEA trial registry.

To guard against over-rejection of the null hypothesis due to multiple inference, we construct a standardized index for each of the seven domains listed above. We construct each index using a weighted average of the standardized index components, where the weights are determined by the inverse covariance matrix of the components (Casey, Glennerster, & Miguel, 2012). The analysis plan describing the exact regression methods to analyze program impacts, along with the details regarding the definitions and construction of the primary and secondary outcome indices, can be found in *Pre-Analysis Plan for Girl Empower* under *Supplementary Materials*, which is also included under the Analysis Plan section of the AEA RCT registry.

Analysis

We compared the baseline values of the seven indices plus age between the three study arms to assess balance at baseline (Appendix Table 1). For each variable, we report the mean in the control group, the differences in GE and GE+ with the control group, the p-value of the difference, and the number of observations, using a linear regression (OLS) model with standard errors clustered at the village level. We also report a "Chi-Squared Test for the Joint Orthogonality of all eight baseline covariates," using a multinomial logit regression with treatment status as the dependent variable and the baseline covariates as the independent variables. To assess whether loss to follow-up rates and the characteristics of those lost to follow-up were similar across study arms, we used the following regression model:

$$Y_{ij} = \alpha + \beta X_{ij} + \gamma^2 T_j^2 + \gamma^3 T_j^3 + \delta^2 T_j^2 X_{ij} + \delta^3 T_j^3 X_{ij} + Z_j + \epsilon_{ij},$$

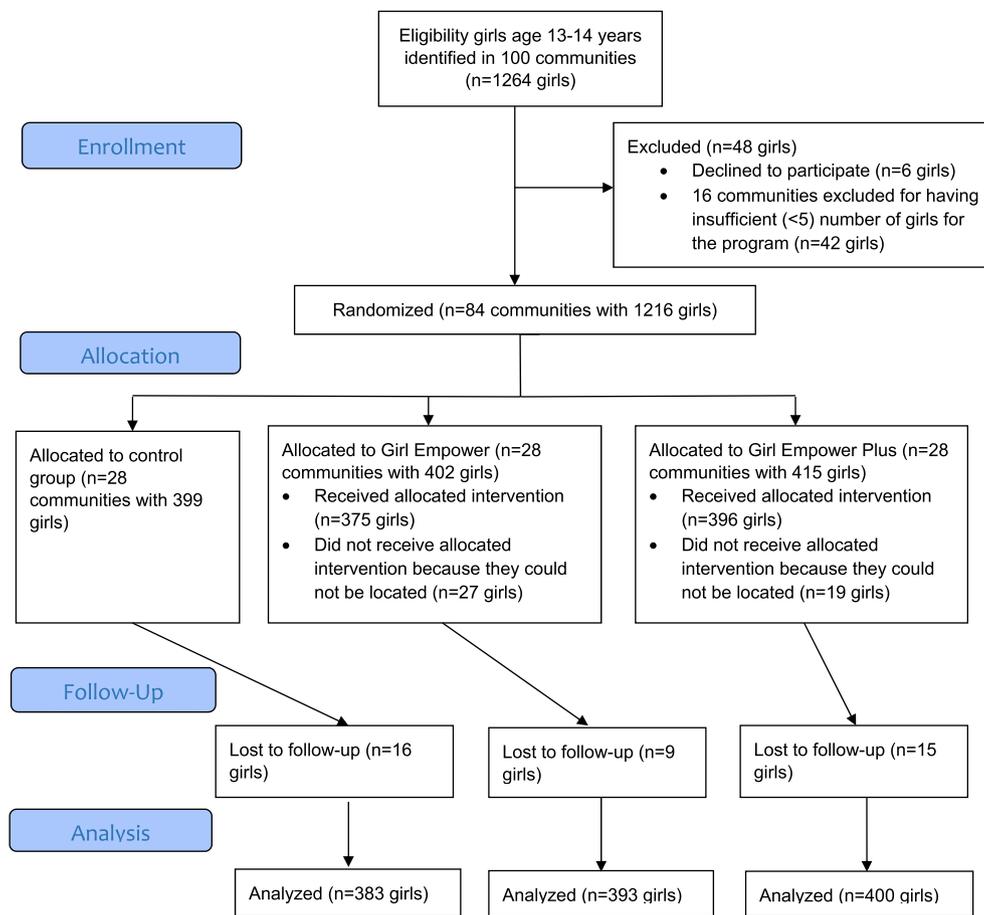


Fig. 1. Study flow diagram.

Table 1
Impacts on sexual and physical violence - 24-month follow-up.

	Dependent Variable:											
	Sexual Violence Index Components										Violence Domain non-component	
	Experienced:										Physical Violence	
	Sexual Violence Index		Non-consensual Touching		Attempted Rape		Pressured Sex		Rape		Physical Violence	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
GE	-0.078 (0.067)	-0.069 (0.069)	0.034 (0.024)	0.038 (0.024)	-0.028 (0.036)	-0.031 (0.035)	0.022 (0.040)	0.011 (0.039)	0.051 (0.040)	0.045 (0.041)	-0.019 (0.017)	-0.019 (0.017)
GE+ (GE + cash)	-0.050 (0.064)	-0.031 (0.060)	0.050** (0.020)	0.046** (0.021)	0.003 (0.035)	-0.002 (0.032)	0.011 (0.034)	0.002 (0.033)	-0.027 (0.034)	-0.028 (0.034)	0.016 (0.013)	0.016 (0.013)
Any Treatment (GE or GE+)	-0.065 (0.054)	-0.050 (0.053)	0.040* (0.021)	0.039* (0.022)	-0.013 (0.028)	-0.017 (0.027)	0.018 (0.033)	0.008 (0.032)	0.015 (0.033)	0.012 (0.033)	-0.003 (0.015)	-0.003 (0.015)
Mean and Standard Deviation of unstandardized dependent variable in the control group	0.000 (1.000)		0.851 (0.357)		0.662 (0.474)		0.455 (0.499)		0.330 (0.471)		0.948 (0.223)	
F-test for Equality of Parameters (p-value)	GE = 0.709	GE+ 0.613	0.471	0.692	0.483	0.503	0.772	0.795	0.064	0.094	0.023	0.020
Lagged Dependent Variable/Age	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	1,175	1,175	1,175	1,175	1,175	1,175	1,175	1,175	1,175	1,175	1,175	1,175

Notes: 0.01 - ***; 0.05 - **, 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. The sexual violence index has been constructed such that higher values imply better outcomes. All index components above are indicator variables. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table.

where Y_{ij} is an indicator for being lost to follow-up for individual i in cluster j , T_j^2 and T_j^3 are binary indicators for GE and GE+, X_{ij} is the vector of eight centered (or demeaned) baseline covariates reported in Appendix Table 1, and Z_j are stratum fixed effects. The findings, along with the mean rate of lost to follow-up in the control group at the 24-month follow-up and joint F-tests for baseline covariates and their interactions with each treatment arm are reported in Appendix Table 2.

To assess program impacts on each domain (and its components), we conducted intent-to-treat (ITT) analysis using linear regression of the following form for outcomes measured at the 24-month follow-up:

$$Y_{ij} = \alpha + \gamma^2 T_j^2 + \gamma^3 T_j^3 + \beta X_{ij} + \varepsilon_{ij},$$

where Y_{ij} is an outcome variable for individual i in cluster j ; T_j^2 and T_j^3 are binary indicators for cluster-level interventions GE and GE+, respectively; and X_{ij} is a vector of baseline covariates consisting only of the lagged dependent variable and age of the individual in years. This is equivalent to the well-known Analysis of Covariance (or ANCOVA) estimation with additional covariate adjustment for age – when the baseline value of the dependent variable is available. The regressions also absorb the strata used for random assignment. The standard errors ε_{ij} , clustered at the village level, account for both the design effect of the cluster-level treatment and heteroskedasticity inherent in the regression model. We report coefficient estimates and p values in tables reporting impacts (Tables 1–7) and in the text when discussing specific impact findings. In the text under the Results section, we also report false discovery rate-adjusted q values for each pairwise comparison of the seven outcome indices between study arms (Appendix Table 4), which is the smallest level of statistical significance at which the null hypothesis is rejected – using the Benjamini and Hochberg (1995) (Benjamini & Hochberg, 1995) procedure as described in Anderson. (Anderson, 2008)

Results

Description of the participants

The baseline characteristics of the study sample have been described elsewhere (Hallman et al., 2016), so we provide only a brief description here. The average age in the control group was 13.7 at baseline. Most

girls (85.03%) had been enrolled in school during the 2014/15 year, before the Ebola outbreak. Overall, 20.7% of girls reported that they had had sex and, of those, 71.4% reported that their first sexual encounter was consensual. Over one-third (37.3%) reported ever experiencing sexual violence: 7.8% had been physically forced to have sex, 8.4% had been non-physically pressured (coerced/persuaded) to have sex, 24.7% had had someone unsuccessfully attempt to have sex with them, and 28.9% had been touched in a sexual way. Less than 2% of the study participants had ever been pregnant or married/cohabitated with a partner. These characteristics were balanced between the control group and the two treatment arms (Hallman et al., 2016).

Baseline balance, loss to follow-up, and participation in Girl Empower.

The vector of baseline characteristics that includes age and the seven outcome indices is very similar across the three study arms, with none of the chi-squared tests for the joint orthogonality of all eight baseline variables producing a p value less than 0.425 (Appendix Table 1). To avoid having an ad hoc vector of background characteristics being included in the baseline balance table, we pre-specified only lagged values of the outcome indicators and age for inclusion, as the latter can be prognostic of many of the outcome indicators, such as sexual debut, pregnancy, and marriage during adolescence.

The rate of loss-to-follow-up was small at 4% in the control group (Appendix Table 2). The attrition analysis utilizes a regression model with an indicator for lost-to-follow-up as the dependent variable and the two treatment indicators along with the baseline characteristics included in Appendix Table 1, which are fully interacted with each treatment indicator, as independent variables. The coefficient estimates for GE and GE+ in column (1) show that the rate of loss to follow-up was similar to the control group in either intervention arm. The joint F-tests of interactions at the bottom of the table under column (2) indicate that the baseline characteristics of those lost to follow-up were not different between GE and GE+.

Attendance in Girl Empower sessions was high in both GE and GE+, with the average number of sessions attended being 27.8 and 28.4, respectively, out of a total of 32 sessions (Appendix Table 3). Caregivers in the GE + arm attended a higher number of monthly caregiver sessions than those in GE (7.45 vs. 6.80; p value < 0.01).

Table 2
Impacts on schooling - 24-month follow-up.

	Dependent Variable:					
	Schooling Index		Schooling Index Components			
	(1)	(2)	(3)	(4)	(5)	(6)
GE	0.049 (0.065)	0.054 (0.070)	0.076 (0.178)	0.162 (0.119)	0.016 (0.021)	0.019 (0.021)
GE+ (GE + cash)	0.065 (0.062)	0.054 (0.057)	-0.032 (0.183)	0.068 (0.122)	0.043* (0.023)	0.044* (0.023)
Any Treatment (GE or GE+)	0.058 (0.054)	0.052 (0.054)	0.030 (0.158)	0.114 (0.110)	0.029 (0.018)	0.032* (0.018)
Mean and Standard Deviation of un standardized dependent variable in the control group	0.000 (1.000)		7.162 (2.205)		0.841 (0.366)	
F-test for Equality of Parameters (p-value) GE = GE+	0.806	1.000	0.526	0.372	0.286	0.322
Lagged Dependent Variable/Age	No	Yes	No	Yes	No	Yes
Number of observations	1,175	1,175	1,175	1,175	1,175	1,175

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. All indices have been constructed such that higher values imply better outcomes. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table. The covariate adjustment referred to as the 'lagged dependent variable' is the baseline value of the dependent variable in all cases, except for attendance, where we control for the baseline value of the schooling index.

Table 3
Impacts on sexual and reproductive health - 24-month follow-up.

	Dependent Variable:											
	Sexual and Reproductive Health Index		Components of Sexual and Reproductive Health Index									
			Never Married		Never Had Sex		Never Pregnant - Extensive Margin		Number of Partners in Past 12 Months - Extensive Margin		Safe Sex Index - Standardized - Extensive Margin	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
GE	0.229*** (0.082)	0.244*** (0.078)	0.024 (0.015)	0.028* (0.015)	0.030 (0.038)	0.036 (0.034)	0.011 (0.024)	0.019 (0.022)	-0.412* (0.231)	-0.458** (0.225)	0.215** (0.089)	0.198** (0.084)
GE+ (GE + cash)	0.354*** (0.087)	0.372*** (0.084)	0.038** (0.014)	0.041*** (0.014)	0.062 (0.040)	0.072** (0.035)	0.018 (0.027)	0.026 (0.026)	-0.651*** (0.203)	-0.705*** (0.203)	0.311*** (0.113)	0.311*** (0.111)
Any Treatment (GE or GE+)	0.291*** (0.077)	0.309*** (0.074)	0.032** (0.013)	0.036** (0.014)	0.046 (0.035)	0.055* (0.032)	0.013 (0.022)	0.022 (0.021)	-0.528*** (0.194)	-0.580*** (0.193)	0.262*** (0.087)	0.255*** (0.085)
Mean and Standard Deviation of unstandardized dependent variable in the control group	0.000 (1.000)		0.922 (0.269)		0.507 (0.501)		0.833 (0.374)		1.364 (3.490)		0.000 (1.000)	
F-test for Equality of Parameters (p-value)	0.085	0.075	0.275	0.281	0.361	0.218	0.803	0.778	0.242	0.198	0.356	0.266
Lagged Dependent Variable/Age	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	1,174	1,174	1,174	1,174	1,174	1,174	1,174	1,174	1,174	1,174	1,174	1,174

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. Both the sexual and reproductive health index and the safe sex index have been constructed such that higher values imply better outcomes. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table. Questions on pregnancy, number of partners in the past 12 months and condom use were only asked to those who reported having ever had sex. The outcome indicators reported above are constructed on the extensive margin - with respondents who reported having never had sex assumed to be never pregnant, having no sexual partners, and assigned the highest score in the safe sex index. The number of partners in the past 12 months questions was administered using the random response method, in which the respondent rolls a die and adds the die roll to the answer and reports the sum to the enumerator. This conceals the true answer from the enumerator who cannot see the roll of the die.

Table 4
Impacts on psychosocial wellbeing - 24-month follow-up.

	Dependent Variable:									
	Psychosocial Index		Psychosocial Wellbeing Index Components						Psychosocial Wellbeing Index non-component	
			Rosenberg Scale		SMFQ Scale		CRIES8 Scale		Suffering from PTSD (CRIES8 < 23)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(11)	(12)	
GE	0.109 (0.069)	0.113 (0.072)	0.035 (0.241)	0.017 (0.240)	0.570 (0.419)	0.645 (0.398)	0.898 (0.578)	1.020* (0.593)	-0.032 (0.033)	-0.036 (0.033)
GE+ (GE + cash)	0.098 (0.070)	0.102 (0.071)	0.191 (0.246)	0.179 (0.248)	0.606 (0.427)	0.530 (0.423)	0.003 (0.635)	0.244 (0.612)	-0.021 (0.036)	-0.028 (0.035)
Any Treatment (GE or GE+)	0.105 (0.065)	0.108 (0.067)	0.116 (0.226)	0.097 (0.224)	0.618 (0.396)	0.610 (0.383)	0.410 (0.507)	0.554 (0.507)	-0.024 (0.028)	-0.029 (0.028)
Mean and Standard Deviation of dependent variable in the control group	0.000 (1.000)		18.399 (2.967)		16.193 (5.913)		18.738 (9.633)		0.754 (0.431)	
F-test for Equality of Parameters (p-value)	0.842	0.853	0.436	0.428	0.925	0.771	0.198	0.250	0.773	0.840
Lagged Dependent Variable/Age	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	1,159	1,159	1,159	1,159	1,159	1,159	1,159	1,159	1,159	1,159

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. All indices have been constructed such that higher values imply better outcomes. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table. The Rosenberg Scale is composed of 10 questions each of which is scored on a Likert scale of 0–3 points, for a total score between 0 and 30 where the higher the score the better self esteem. The SMFQ (Short Mood of Feelings Questionnaire) is composed of 12 questions each of which is scored on a Likert scale of 0–2 points, for a total score between 0 and 24 where higher values imply lower depression. CRIES8 (Children’s Revised Impact of Events) scale is composed of 8 questions scored on an inverted Likert scale of 0–5 points, for a total score between 0 and 40, so that higher scores indicate lower levels of trauma. The scale was administered to respondents who experienced sexual or physical violence during the 24-month period between baseline and endline data collection (95.3%). Those that did not experience any violence were assigned the maximum score of 40. Respondents with a score below 23 on our inverted scale (equivalent to above 17 in the original scale), which was used as a referral value for IRC, are considered as suffering from PTSD.

Table 5
Impacts on gender attitudes - 24-month-Follow up.

	Dependent Variable:					
	Gender Attitudes Index Components (Standardized)					
	Gender Equity Index		Attitudes Towards IPV Index			
	(1)	(2)	(3)	(4)	(5)	(6)
GE	0.207**	0.206**	0.139	0.141*	0.181**	0.177*
	(0.088)	(0.088)	(0.085)	(0.083)	(0.090)	(0.090)
GE+ (GE + cash)	0.224***	0.228***	0.133	0.137*	0.213***	0.215***
	(0.082)	(0.081)	(0.084)	(0.081)	(0.074)	(0.075)
Any Treatment (GE or GE+)	0.216***	0.215***	0.142*	0.142*	0.191**	0.191**
	(0.077)	(0.076)	(0.078)	(0.077)	(0.075)	(0.075)
Mean and Standard Deviation of un standardized dependent variable in the control group	0.000		0.005		-0.001	
	(1.000)		(0.995)		(1.001)	
F-test for Equality of Parameters (p-value) GE = GE+	0.823	0.773	0.932	0.954	0.688	0.640
Lagged Dependent Variable & Age	No	Yes	No	Yes	No	Yes
Number of observations	1,161	1,161	1,161	1,161	1,161	1,161

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. All indices have been constructed such that higher values imply better outcomes. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table.

Table 6
Impacts on life skills - 24-month-Follow up.

	Dependent Variable:											
	Life Skills Index Components (Standardized)											
	Life Skills Index		Knowledge of HIV/AIDS		Health		Financial Literacy		Knowledge of Condom Effectiveness		Healthy Intimate (Heterosexual) Relationships	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
GE	0.233**	0.224**	0.198**	0.193*	0.129	0.109	0.216***	0.220***	0.191**	0.177**	-0.025	-0.023
	(0.098)	(0.100)	(0.096)	(0.103)	(0.088)	(0.087)	(0.072)	(0.073)	(0.078)	(0.078)	(0.112)	(0.111)
GE+ (GE + cash)	0.294***	0.289***	0.166**	0.142*	0.076	0.076	0.342***	0.342***	0.261***	0.249***	-0.030	-0.036
	(0.096)	(0.094)	(0.081)	(0.083)	(0.084)	(0.084)	(0.080)	(0.080)	(0.072)	(0.071)	(0.112)	(0.113)
Baseline Lagged Value Missing > 10% = 1					0.154	0.154						
					(0.160)	(0.160)						
Any Treatment (GE or GE+)	0.260***	0.251***	0.176**	0.157*	0.099	0.092	0.280***	0.280***	0.222***	0.212***	-0.029	-0.025
	(0.086)	(0.086)	(0.079)	(0.083)	(0.077)	(0.075)	(0.065)	(0.066)	(0.071)	(0.070)	(0.102)	(0.100)
Mean and Standard Deviation of un standardized dependent variable in the control group	0.000		0.022		-0.001		0.004		0.027		0.011	
	(1.000)		(0.987)		(1.001)		(1.004)		(0.987)		(1.005)	
F-test for Equality of Parameters (p-value) GE = GE+	0.508	0.478	0.717	0.568	0.540	0.699	0.121	0.128	0.186	0.170	0.956	0.894
Lagged Dependent Variable/ Age	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156	1,156

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. All indices have been constructed such that higher values imply better outcomes. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 individuals from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table.

Program impacts

We assessed the causal effects of GE and GE + on seven domains (Fig. 2), which are presented in greater detail in Tables 1–7. Our primary outcome was sexual violence, the self-reported rates of which are very high in the study population: for example, 33.0% reported having been raped during the 24-month period between baseline and endline data collection (Table 1). Almost everyone (94.8%) reported having experienced physical violence during the same period. Neither GE nor GE +

had a statistically significant effect on the sexual violence index (or on physical violence). The standardized effects of either program on the schooling index were small ($\beta \leq 0.07$) and not statistically significant at the 95% level of confidence (Table 2). 16% of the study population of 15-16-year-old females at endline were not enrolled in school during the 2016-17 school year. Both programs caused significant improvements in the standardized SRH index; the coefficient estimate of the ITT effect was more than 50% higher in GE + than GE (GE: 0.244 SDs, $p < 0.01$; GE+: 0.372 SDs, $p < 0.01$; F-test for GE = GE+: $p = 0.075$). This benefit

Table 7
Impacts on protective factors - 24 month-follow-up.

	Dependent Variable:							
	Protective Factors Index		Protective Factors Index Components (Standardized)					
			Social Capital Index		Gender Norms Index (Caregiver Survey)		Child Rearing Index (Caregiver Survey)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GE	0.012 (0.105)	0.019 (0.101)	0.025 (0.100)	0.029 (0.100)	0.015 (0.110)	0.008 (0.110)	-0.025 (0.085)	-0.010 (0.086)
GE+ (GE + cash)	0.126 (0.107)	0.099 (0.106)	0.134 (0.098)	0.123 (0.098)	0.073 (0.104)	0.038 (0.106)	-0.000 (0.094)	-0.011 (0.095)
Any Treatment (GE or GE+)	0.070 (0.098)	0.064 (0.094)	0.086 (0.094)	0.085 (0.092)	0.040 (0.101)	0.019 (0.102)	-0.015 (0.080)	-0.008 (0.080)
Mean and Standard Deviation of dependent variable in the control group	0.000 (1.000)		0.004 (0.985)		-0.006 (1.007)		0.027 (0.994)	
F-test for Equality of Parameters (p-value) GE = GE+	0.249	0.421	0.179	0.264	0.491	0.727	0.780	0.992
Lagged Dependent Variable/Age	No	Yes	No	Yes	No	Yes	No	Yes
Number of observations	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052

Notes: 0.01 - ***; 0.05 - **; 0.1 - *; OLS regressions at the individual (adolescent female respondent) level at the 24-month follow-up with standard errors (SEs) in parentheses clustered at the village level. All regressions include stratum fixed effects. All indices have been constructed such that higher values imply better outcomes. The gender norms and child rearing indices are based on questions administered to the primary caregiver of the adolescent female respondent. Please refer to the pre-analysis plan in the appendix for index creation and the construction of index sub-components. 40 adolescent female respondents from baseline were lost to the 24-month follow-up, leaving 1,176 individuals who were interviewed at endline. Of these 1,176 individuals, the primary caregivers of 1,078 were successfully interviewed at the 24-month follow-up. The sample is limited to only those observations with no item non-response for any of the index components or sub-components. Robustness of impact findings to how item non-response was handled is presented in Appendix A. Missing baseline covariates were replaced with the sample mean. If more than 10% of a baseline covariate was missing, an indicator variable for missing baseline covariate was also included in the regression and reported in the table.

was apparent in four of the five index components – never married, never had sex, number of sexual partners in the past 12 months, and the safe sex index (Table 3). The safe sex index includes questions on the frequency of condom use during the past 12 months and whether a condom was used the last time the respondent had sex. The effects on all index components are meaningfully larger in GE + than GE, but the study is only powered to detect differences in the overall index between GE and GE+, and not for individual index components.

We also assessed program impacts on the health, gender attitudes, and knowledge of the study participants. Both variants of GE raised the psychosocial index by about 0.1 SD, but these effects were not statistically significant. Three quarters of the study population suffered elevated levels of distress, based on their Children’s Revised Impact of Event Scale (CRIES8) scores – a PTSD scale (Table 4). Consenting participants with high CRIES8 scores received a visit from an IRC social worker to provide them with counseling and referrals to health service providers. Both programs had significantly higher standardized index scores on Gender Attitudes compared to those in the control group (GE: 0.206 SD, p<0.05; GE+: 0.228 SD, p<0.05). This benefit was seen for both components of the index – gender equity and attitudes towards IPV – with no statistically significant differences in impacts between GE and GE+ (Table 5). We also found that girls in either program had higher standardized index scores on Life Skills (GE: 0.224 SDs, p<0.05; GE+: 0.289 SDs, p<0.01) and the benefit was apparent in three of the five components of the index – knowledge of HIV, financial literacy, and knowledge of condom effectiveness (Table 6). Finally, neither variant of GE had a detectable effect on an index of protective factors for the beneficiaries, which include their social capital, the attitudes of their primary caregivers towards gender equity and their aspirations for their daughters in terms of their education, age at first marriage and pregnancy, and labor market participation (Table 7).

Adjustments for multiple hypothesis testing, i.e. for the likelihood of finding statistically significant treatment effects on the seven indices when comparing the combined GE program to the control group by chance, did not alter the main takeaways from Tables 1–7 discussed above. For the three domains on which the program had statistically significant effects, the False Discovery Rate-adjusted (FDR) q-values were all below 0.05 (SRH: FDR q = 0.004, Gender Attitudes: FDR q = 0.014, Life Skills: FDR q = 0.014).

Discussion

The *Girl Empower* program was not successful in attaining its primary goal, which was to reduce the incidence of sexual violence experienced by adolescent females. As the program did not increase the protective factors surrounding adolescent girls – the social network of program beneficiaries was not affected, nor were the caregivers’ gender attitudes or their aspirations for the girl children – it is perhaps not surprising that the incidence of sexual violence experienced by program beneficiaries did not decline (Yount et al., 2017). While Liberia has signed international laws that protect women and girls, customary laws like early marriage are still practiced in some communities, while at the same time forbidding discussing sexual education openly or with children. It is also possible that within the combined post-Ebola and post-conflict environment, factors discussed by Kelly et al. (2018) (Kelly et al., 2018), such as greater habituation to conflict, weakening of social support structures, shifts in family roles and responsibilities, and disruption of peaceful conflict resolution practices, prevented the hypothesized impact. The *Girl Empower* program engaged with girls and parents, but it did not have a systematic strategy to engage with community structures on these topics. In addition, while *Girl Empower* trained service providers on gender-based violence response for adolescent girls, there were few service providers in the area, often far from *Girl Empower* implementation sites. The dearth of services also includes mental health and psychosocial support for adolescent girls who experienced violence, in order to help them heal and recover as well as cope with continued risks of violence in their homes, schools, communities, play grounds, and market places. While program effects on schooling and psychological wellbeing were positive, the standardized effect sizes were small (approximately 0.1 SD or less) and statistically non-significant. The small effects on education (statistically non-significant in GE and only significant at the 90% level of confidence in GE+ in Table 2) are consistent with evaluations of similar “safe spaces” programs for adolescent females (Bandiera et al., forthcoming; Buchmann et al., 2018).

Both programs did, however, seem to have moderate and statistically significant effects on three domains: Gender Attitudes, Life Skills, and SRH. The effects of both GE and GE+ were above 0.2 SD for all three indices, and as large as 0.37 SD for the SRH index in GE+. Beneficiaries

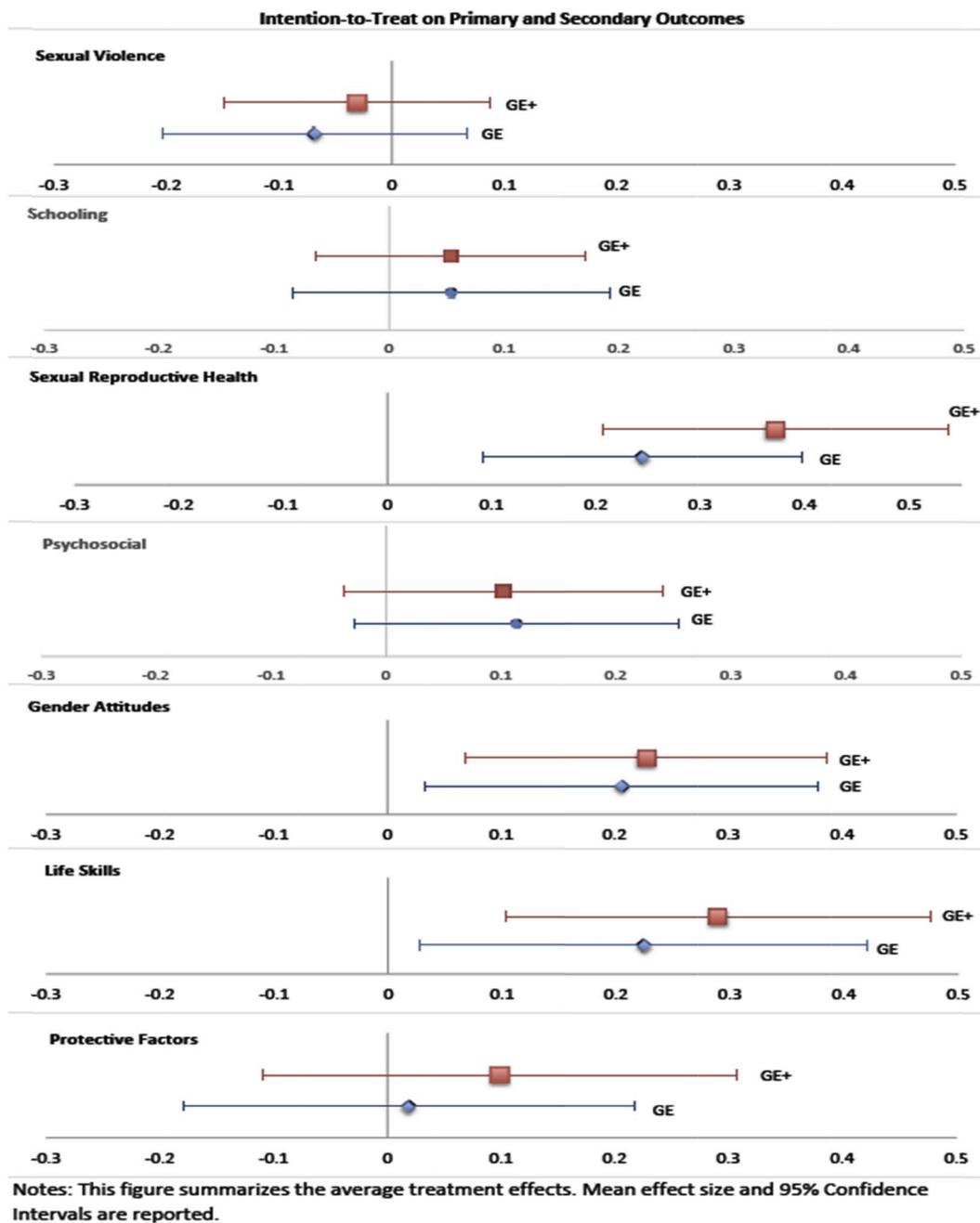


Fig. 2. Summary of ITT effects in GE and GE+ in comparison with the control group

of both programs were less likely to be accepting of IPV and they showed a better understanding of condom effectiveness, HIV/AIDS, and financial matters, such as saving, budgeting, borrowing, etc. They were also 3.6 percentage points (or 46%) less likely to be married, had a lower number of sexual partners (0.58 fewer partners, 43% lower), and were more likely to practice safer sex compared with the control group.

GE and GE + produced similar impacts in all but one domain. The effect size on the SRH index, as well as each of its components, was approximately 50% higher in GE + than GE. GE + reduced the likelihood of marriage and the number of sexual partners in the past 12 months by more than 50%. Our ex ante hypothesis as to why GE + may be more effective than GE was that the small cash incentives for attendance would increase program participation in the target population. Attendance, however, was similarly high in both GE and GE+. While the cash incentives increased the likelihood of caregivers attending the monthly sessions, we do not think that the modest increase in caregiver

attendance (6.80 vs. 7.45 out of a total of eight sessions) can explain the differential effects on SRH. As mentioned above, the program had no effect on caregivers' gender attitudes or their aspirations for the girl children. So, what explains the higher impacts on SRH in GE + than GE?

A potential mechanism is a pure income effect. Previous studies have shown that positive income shocks can reduce early marriages and teen pregnancies in both developing (Baird et al., 2011, 2012; Robinson & Yeh, 2011) and developed countries (Hankins & Hoekstra, 2011). Study participants in both GE and GE + received the US\$2/month savings payments, for a modest sum of US\$16 during the *Girl Empower* program, which were provided unconditionally, i.e. not tied to program participation. While the original study design envisioned the participation incentives to be small (US\$0.50 per session attended) in order to minimize the possibility of a direct income effect on the outcomes of interest, the actual incentive payment was US\$1.25 per session attended – due to a miscommunication between the study team and the IRC Liberia staff.

As mentioned earlier, the implied transfer to the caregiver of a regular program participant of approximately US\$6/month constitutes more than 10% of per capita consumption in Liberia (World Bank, 2018), and likely more in the study area, because Nimba is a poorer than average rural county in Liberia. As such, it represents a non-negligible increase to household income during the 8-month program, which could have directly affected child marriage and sexual behavior among GE + beneficiaries. However, the fact that the 24-month follow-up data was collected approximately one year after the last cash transfer payment also raises the possibility of an indirect income effect: perhaps, the additional income provided the space for the GE + participants to better internalize the lessons from the mentoring program and to reinforce their newly obtained knowledge and skills towards sustained behavior change. All participating caregivers were engaged in sessions on the effects of sexual abuse and early marriage on adolescent girls, their families, and the community at large. While the caregivers in the GE + groups were free to use the cash transfers as they saw fit, they were encouraged to use these funds in ways that would promote the adolescent girls' safety and well-being. In this scenario, a positive interaction effect between the mentoring program and cash transfers would produce a larger effect than either mentoring or cash transfers alone. However, since we did not include a study arm providing cash transfers alone, we cannot formally test this hypothesis.

Other trials that included a similar safe spaces, mentoring, or empowerment intervention targeted to adolescent females do not seem particularly helpful in providing further evidence on the relative roles of program components. A trial that contrasted an empowerment program with a incentives (in the form of cooking oil to the parents) conditional on the adolescent female not getting married until the age of 18 found that the conditional transfer program was effective in reducing child marriage and teen fertility, but the empowerment program was not. Both programs had small effects on school participation and attainment and there were no complementarities between the two (Buchmann et al., 2018). The nature of the in-kind transfer (conditional on not getting married as a child) is quite different than complementing the empowerment program with small transfers conditional on attendance, limiting the applicability of those findings to this study in Liberia. Another trial evaluating the effects of providing vocational and life skills through girls' clubs in Uganda found that it was highly effective in reducing early marriage, teen pregnancy, and nonconsensual sexual activity (Bandiera et al., forthcoming). However, the authors are unable to comment on the relative importance of vocational vs. life skills components of the intervention (or that of simply having access to safe spaces to gather and/or have access to an older mentor), but they do emphasize the multi-faceted nature of the intervention as a possible factor behind its success.

The study had some limitations. At baseline all interview questions were face-to-face verbal. The same procedure was used at endline with the exception of questions pertaining to violence, which were pre-recorded in every local language so that the respondents could listen to an audio recording of each question in the language of their choice (with the tablet held to the ear, as if listening to a phone call), after which their response was indicated to the female enumerator who entered it in the tablet. This was done to ensure that the respondents understood each question exactly, as the definitions for various types of sexual violence experienced could be detailed and complex. While the study was well powered to detect standardized effects of 0.1 to 0.2 SD for any of the seven main indices in pairwise comparisons between any of the three study arms, it was not well powered to detect impacts on individual index sub-components. Primary and secondary outcome indices were constructed using self-reported data, which could have caused spurious findings if there was differential misreporting between study arms.

Girl Empower aimed to endow adolescent girls with life and financial skills, expand their support networks, and provide them with safe spaces to discuss topics relevant to their daily lives. The study showed that *Girl*

Empower could equip adolescent females with important life skills, positively influence their gender attitudes, and, perhaps most importantly, improve their sexual and reproductive health, all of which were sustained 12 months after the end of the program. Small incentive payments for regular attendance were effective in enhancing the impact of the program on SRH. Future trials may experiment with comparing *Girl Empower* with cost-equivalent cash transfers, in addition to *Girl Empower* plus cash as we have done here. It is also important, however, to identify how programs like *Girl Empower* can be modified or enhanced to reduce the sexual abuse of adolescent girls, the rates of which are unacceptably high, especially in conflict and post-conflict settings like the one studied here.

Declaration of competing interest

None of the authors have any conflicts of interest to disclose.

Acknowledgments

The authors would like to thank the research team at Innovations for Poverty Action Liberia for leading data collection efforts; Chiara Dall'Aglio for capable research assistance; the Liberian Ministry of Gender, Children and Social Protection and the Coalition Of Women Against Gender Base Violence in Liberia (COGAWIL) for their collaboration; the IRC Liberia country program, the IRC Violence Prevention and Response Unit, Dr. Juliette Seban, and Dr. Kathryn Falb; and finally all the communities, adolescent girls, and caregivers who took part in the study. Hallman, Özler, and Guimond presented the findings reported in the manuscript in an oral presentation at Interaction in Washington, D.C. and at the International Rescue Committee in New York City in April 2018.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2019.100527>.

Funding source

NoVo Foundation funded this research. NoVo Foundation had no role in the writing of the manuscript or the decision to submit it for publication.

References

- Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484), 1481–1495.
- Austrian, K., & Muthengi, E. (2014). Can economic assets increase girls' risk of sexual harassment? Evaluation results from a social, health and economic asset-building intervention for vulnerable adolescent girls in Uganda. *Children and Youth Services Review*, 47, 168–175.
- Baird, S. J., Garfein, R. S., McIntosh, C. T., & Ozler, B. (2012). Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: A cluster randomised trial. *Lancet*, 379(9823), 1320–1329.
- Baird, S., McIntosh, C. T., & Ozler, B. (2011). Cash or condition? Evidence from a cash transfer experiment. *Quarterly Journal of Economics*, 126(4), 1709–1753.
- Baird, S., McIntosh, C., & Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics*, 140, 169–185.
- Bandiera, O., Buehren, N., Burgess, R., et al. (2019). Women's empowerment in action: Evidence from a randomized control trial in africa. *American Economic Journal: Applied Economics* (forthcoming).
- 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(1), 289–300.
- Bjorkman Nyqvist, M., Corno, L., de Walque, D., & Svensson, J. (2018). Incentivizing safer sexual behavior: Evidence from a lottery experiment on HIV prevention. *American Economic Journal: Applied Economics*, 10(3), 287–314.
- Bruce, J., & Hallman, K. (2008). Reaching the girls left behind. *Gender and Development*, 16(2), 227–245.

- Buchmann, N., Field, E., Glennerster, R., Nazneen, S., Pimkina, S., & Sen, I. (2018). In *Power vs money: Alternative approaches to reducing child marriage in Bangladesh, a randomized control trial*. <https://www.povertyactionlab.org/sites/default/files/publications/Power-vs-Money-Working-Paper.pdf>. (Accessed 14 October 2019).
- Buller, A. M., Hidrobo, M., Peterman, A., & Heise, L. (2016). The way to a man's heart is through his stomach?: A mixed methods study on causal mechanisms through which cash and in-kind food transfers decreased intimate partner violence. *BMC Public Health*, 16, 488.
- Buller, A., Peterman, A., Ranganathan, M., Bleile, A., Hidrobo, M., & Heise, L. (2018). In *A mixed-method review of cash transfers and intimate partner violence in low and middle-income countries*. <https://doi.org/10.18356/f5ce8f94-en>. (Accessed 3 January 2019).
- Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *Quarterly Journal of Economics*, 127(4), 1755–1812.
- Cluver, L., Boyes, M., Orkin, M., Pantelic, M., Molwena, T., & Sherr, L. (2013). Child-focused state cash transfers and adolescent risk of HIV infection in South Africa: A propensity-score-matched case-control study. *The Lancet Global health*, 1(6), e362–e370.
- Cluver, L. D., Orkin, F. M., Boyes, M. E., & Sherr, L. (2014). Cash plus care: Social protection cumulatively mitigates HIV-risk behaviour among adolescents in South Africa. *AIDS*, 28(Suppl 3), S389–S397.
- Duflo, E., Dupas, P., & Kremer, M. (2015). Education, HIV, and early fertility: Experimental evidence from Kenya. *The American Economic Review*, 105(9), 2757–2797.
- Dunbar, M. S., Kang Dufour, M. S., Lambdin, B., Mudekunya-Mahaka, I., Nhamo, D., & Padian, N. S. (2014). The SHAZI Project: Results from a pilot randomized trial of a structural intervention to prevent HIV among adolescent women in Zimbabwe. *PLoS One*, 9(11), e113621.
- Dunbar, M. S., Maternowska, M. C., Kang, M. S., Laver, S. M., Mudekunya-Mahaka, I., & Padian, N. S. (2010). Findings from SHAZI: A feasibility study of a microcredit and life-skills HIV prevention intervention to reduce risk among adolescent female orphans in Zimbabwe. *Journal of Prevention & Intervention in the Community*, 38(2), 147–161.
- Dziewanski, D. (2012). In *Peace without security: Violence against women and girls in Liberia*. <http://www.smallarmssurvey.org/fileadmin/docs/G-Issue-briefs/Liberia-AV-A-IB3.pdf>. (Accessed 7 December 2018).
- Ellsberg, M., Vyas, A., Madrid, B., Quintanilla, M., Zelaya, J., & Stockl, H. (2017). Violence against adolescent girls: Falling through the cracks? *Background paper*. https://globalwomensinstitute.gwu.edu/sites/g/files/zaxdzs1356/f/downloads/Falling%20through%20the%20Cracks_Background%20Paper%20%281%29.pdf. (Accessed 14 October 2019).
- Gibbs, A., Corboz, J., Shafiq, M., et al. (2018). An individually randomized controlled trial to determine the effectiveness of the women for women international programme in reducing intimate partner violence and strengthening livelihoods amongst women in Afghanistan: Trial design, methods and baseline findings. *BMC Public Health*, 18(1), 164.
- Gibbs, A., Jacobson, J., & Kerr Wilson, A. (2017). A global comprehensive review of economic interventions to prevent intimate partner violence and HIV risk behaviours. *Global Health Action*, 10(sup2), 1290427.
- Guedes, A., Bott, S., Garcia-Moreno, C., & Colombini, M. (2016). Bridging the gaps: A global review of intersections of violence against women and violence against children. *Global Health Action*, 9, 31516.
- Hallman, K., Kelvin, E., Ozler, B., et al. (2016). In *Combining mentoring programs with cash transfers for adolescent girls in Liberia: Baseline report*. <http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-7797> accessed May 10, 2-18.
- Handa, S., Peterman, A., Huang, C., Halpern, C., Pettifor, A., & Thirumurthy, H. (2015). Impact of the Kenya cash transfer for orphans and vulnerable children on early pregnancy and marriage of adolescent girls. *Social Science & Medicine*, 141, 36–45.
- Hankins, S., & Hoekstra, M. (2011). Lucky in life, unlucky in love? The effect of random income shocks on marriage and divorce. *Journal of Human Resources*, 46(2), 403–426.
- Haushofer, J., Ringdal, C., Shapiro, J., & Wang, X. (2019). In *Income changes and intimate partner violence: Evidence from unconditional cash transfers in Kenya*. <https://www.nber.org/papers/w25627>. (Accessed 14 October 2019).
- Heath, R., Hidrobo, M., & Roy, S. (2018). In *Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali*. <http://ebrary.ifpri.org/cdm/singleitem/collection/p15738coll2/id/133031>. (Accessed 14 October 2019).
- Heinrich, C., HJ, H., & Samson, M. (2017). Reducing adolescent risky behaviors in a high-risk context: The effects of unconditional cash transfers in South Africa. *Economic Development and Cultural Change*, 65(4), 619–652.
- Hidrobo, M., Peterman, A., & Heise, L. (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern Ecuador. *American Economic Journal: Applied Economics*, 8(3), 284–303.
- Jewkes, R., Gibbs, A., Jama-Shai, N., et al. (2014). Stepping stones and creating futures intervention: Shortened interrupted time series evaluation of a behavioural and structural health promotion and violence prevention intervention for young people in informal settlements in durban, South Africa. *BMC Public Health*, 14, 1325.
- Kelly, J. T. D., Colantuoni, E., Robinson, C., & Decker, M. R. (2018). From the battlefield to the bedroom: A multilevel analysis of the links between political conflict and intimate partner violence in Liberia. *BMJ global health*, 3(2), e000668.
- Kilburn, K. N., Pettifor, A., Edwards, J. K., et al. (2018). Conditional cash transfers and the reduction in partner violence for young women: An investigation of causal pathways using evidence from a randomized experiment in South Africa (HPTN 068). *Journal of the International AIDS Society*, 21(Suppl 1).
- Kohler, H. P., & Thornton, R. (2012). Conditional cash transfers and HIV/AIDS prevention: Unconditionally promising? *The World Bank Economic Review*, 26(2), 165–190.
- Liberia Institute of Statistics and Geo-Information Services [LISGIS] Liberia. (2007). In *Ministry of health and social welfare [Liberia], national AIDS control program [Liberia]*. Macro International Inc. Liberia Demographic and Health Survey, 2008 <https://dhsprogram.com/pubs/pdf/fr201/fr201.pdf>. (Accessed 8 June 2018).
- Liberia Institute of Statistics and Geo-Information Services [LISGIS] Liberia. (2014). In *Ministry of health and social welfare [Liberia], national AIDS control program [Liberia]*. International. I. Liberia demographic and health survey 2013. <https://dhsprogram.com/pubs/pdf/FR291/FR291.pdf>. (Accessed 8 June 2018).
- Lundgren, R., & Amin, A. (2015). Addressing intimate partner violence and sexual violence among adolescents: Emerging evidence of effectiveness. *Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, 56(1 Suppl), S42–S50.
- Marcus, R., Gupta-Archer, N., Darcy, M., & Page, E. (2017). In *Girls' clubs, life skills programmes and girls' well-being outcomes*. <https://www.gage.odi.org/sites/default/files/2017-10/GAGE%20Girls%20Club%20Report%20FINAL.pdf>. (Accessed 11 May 2018).
- Minnis, A. M., vanDommelen-Gonzalez, E., Luecke, E., Dow, W., Bautista-Arredondo, S., & Padian, N. S. (2014). Yo puedo—a conditional cash transfer and life skills intervention to promote adolescent sexual health: Results of a randomized feasibility study in san francisco. *Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, 55(1), 85–92.
- Peterman, A., Neijhoft, A. N., Cook, S., & Palermo, T. M. (2017). Understanding the linkages between social safety nets and childhood violence: A review of the evidence from low- and middle-income countries. *Health Policy and Planning*, 32(7), 1049–1071.
- Pettifor, A., MacPhail, C., Hughes, J. P., et al. (2016). The effect of a conditional cash transfer on HIV incidence in young women in rural South Africa (HPTN 068): A phase 3, randomised controlled trial. *The Lancet Global health*, 4(12), e978–e988.
- Robinson, J., & Yeh, E. (2011). Transactional sex as a response to risk in Western Kenya. *American Economic Journal: Applied Economics*, 3, 35–64.
- Rosenberg, M., Pettifor, A., Thirumurthy, H., Halpern, C. T., & Handa, S. (2014). The impact of a national poverty reduction program on the characteristics of sex partners among Kenyan adolescents. *AIDS and Behavior*, 18(2), 311–316.
- Roy, S., Hidrobo, M., Hoddinott, J., & Ahmed, A. (2019). Transfers, behavior change communication, and intimate partner violence: Post program evidence from rural Bangladesh. *The Review of Economics and Statistics*, 101(5), 1–13.
- Swiss, S., Jennings, P. J., Aryee, G. V., et al. (1998). Violence against women during the Liberian civil conflict. *Jama*, 279(8), 625–629.
- Temin, M., Amin, S., Ngo, T., & Psaki, S. (2018). In *How to give adolescent girls voice, choice, and control. A framework for programs focused on empowering adolescent girls*. https://ssir.org/articles/entry/how_to_give_adolescent_girls_voice_choice_and_control. (Accessed 14 October 2019).
- Walque, D., Dow, W. H., Nathan, R., et al. (2012). Incentivising safe sex: A randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ open*, 2, e000747.
- Walque, D., Fernald, L., Gertler, P., & Hidrobo, M. (2017). Cash transfers and child and adolescent development. In S. N. Bundy DAP, S. Horton, et al. (Eds.), *Child and adolescent health and development* (3rd ed.). Washington DC, USA: The International Bank for Reconstruction and Development /The World Bank.
- World Bank. (2018). In *Select country aggregations: Liberia 2011*. <http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx>. (Accessed 11 May 2018).
- Yount, K. M., Krause, K. H., & Miedema, S. S. (2017). Preventing gender-based violence victimization in adolescent girls in lower-income countries: Systematic review of reviews. *Social Science & Medicine*, 192, 1–13.