B For Online Publication – Appendices

B.1 Theoretical Appendices

B.1.1 Proof of proposition one

The couple will choose $s = 1$ if and only if there exists some readily available technology on the frontier (P, H) such that $u_i(P, H) \geq u_i^0 \quad \forall i = m, f$. Define

$$I^0 = \left\{(P, H) \in R^2 | u_i(P, H) \geq u_i^0, i = m, f\right\} \tag{B.1}$$

as the set of all points (P, H) that satisfies both partner's participation constraints. [28]

To see why the optimal choice of health is increasing in α, assume that the intersection $\{US, FC, MC\} \cap I^0$ is non-empty, and thus that sex with some readily available technology provides greater utility to both members of the couple than no sex. Consider then the unconstrained household maximisation problem

$$\max_H \left\{\alpha u_f(P(H), H) + (1 - \alpha) u_m(P(H), H)\right\}. \tag{B.2}$$

Since each $u_i(P(H), H)$ is quasi-concave, the objective function is also quasi-concave and has a unique solution. Denote this solution $\widetilde{H}(\alpha)$. It follows straightforwardly from the single crossing property in Assumption 1 that $\widetilde{H}'(\alpha) > 0$.

For convenience of notation, define

$$U_j(H) = u_j(P(H), H) \tag{B.3}$$

for partner $j = m, f$, where $P(H)$ describes the technological frontier. Equation B.2 becomes

$$\max_H \left\{\alpha U_f(H) + (1 - \alpha) U_m(H)\right\}. \tag{B.4}$$

[28] Specifically, $I^0 = I_m^0 \cap I_f^0$, where $I_i^0 = \left\{(P, H) \in R^2 | u_i(P, H) \geq u_i^0\right\}$ is the upper contour set of the indifference curve corresponding to the reservation utility u_i^0.

The first-order condition is

$$\alpha U_f'(H) + (1-\alpha) U_m'(H) = 0. \tag{B.5}$$

Note this implies that at the optimal choice $\widetilde{H}$, U_f' and U_m' must be of opposite signs. It follows from the single-crossing property in Equation 1 that at the optimum, $U_f'(H) > 0$ and $U_m'(H) < 0$.

The second-order condition is

$$\alpha U_f''(H) + (1-\alpha) U_m''(H) < 0. \tag{B.6}$$

Taking the first-order condition in Equation B.5 as an implicit definition of $\widetilde{H}(\alpha)$, and differentiating with respect to α, we obtain

$$\left[\alpha U_f''(H(\alpha)) + (1-\alpha) U_m''(H(\alpha))\right] \widetilde{H}'(\alpha) + U_f'(H) - U_m'(H) = 0, \tag{B.7}$$

which yields

$$\widetilde{H}'(\alpha) = -\frac{U_f'(H) - U_m'(H)}{\alpha U_f''(H(\alpha)) + (1-\alpha) U_m''(H(\alpha))}. \tag{B.8}$$

To determine the sign of the numerator, note that from the first-order condition we have that

$$-U_m'(H) = \frac{\alpha}{1-\alpha} U_f'(H), \tag{B.9}$$

and thus that

$$sgn\left[\widetilde{H}'(\alpha)\right] = sgn\left[U_f'(H) - U_m'(H)\right] = sgn\left[U_f'(H)\left(1 + \frac{\alpha}{(1-\alpha)}\right)\right] = sgn\left[U_f'(H)\right]. \tag{B.10}$$

As reasoned above, at the optimum $U_f'(H) > 0$ because of the single-crossing property. Thus $\widetilde{H}'(\alpha) > 0$.

However, it is possible that $\widetilde{H}(\alpha)$ does not lie on the intersection of I^0 and the technology frontier. By the single crossing assumption, the left-most endpoint H_L of this intersection is defined by $u_f(P(H_L), H_L) = u_f^0$, while the right-most endpoint H_U

is defined by $u_f\left(P\left(H_U\right), H_U\right) = u_m^0$. This is illustrated in Figure 1. It could therefore be that $u_f\left(P\left(\widetilde{H}\left(\alpha\right)\right), \widetilde{H}\left(\alpha\right)\right) < u_f^0$ or that $u_m\left(P\left(\widetilde{H}\left(\alpha\right)\right), \widetilde{H}\left(\alpha\right)\right) < u_m^0$ (but not both). Consider the case in which her participation constraint binds, such that $u_f\left(P\left(\widetilde{H}\left(\alpha\right)\right), \widetilde{H}\left(\alpha\right)\right) < u_f^0$. The couple then instead chooses the closest incentive-compatible choice, which solves the incentive-constrained household utility maximisation problem

$$\max_H \left\{u_m\left(P(H), H\right) | \mu_f\left[u_f\left(P(H), H\right) - u_f^0\right]\right\}. \tag{B.11}$$

They hence choose H_L, which is independent of α. Vice versa, if his participation constraint binds they choose H_U. If neither partner's participation constraint binds, they choose $\widetilde{H}\left(\alpha\right)$ as before.

Given that $\widetilde{H}\left(\alpha\right)$ is increasing in α, this implies that there are threshold values for α defined by $\widetilde{H}\left(\alpha_j\right) = H_j$ for $j = L, U$ such that

$$H^*\left(\alpha\right) = \begin{cases} H_L & \text{if } \alpha < \alpha_L \\ \widetilde{H}\left(\alpha\right) & \text{if } \alpha \in [\alpha_L, \alpha_U] \\ H_U & \text{if } \alpha > \alpha_U. \end{cases} \tag{B.12}$$

It follows that $H^*\left(\alpha\right)$ is weakly increasing in α: $H^*\left(\alpha\right)$ is constant below α_L and above α_H, and is strictly increasing inbetween. This is illustrated in Figure B.1.

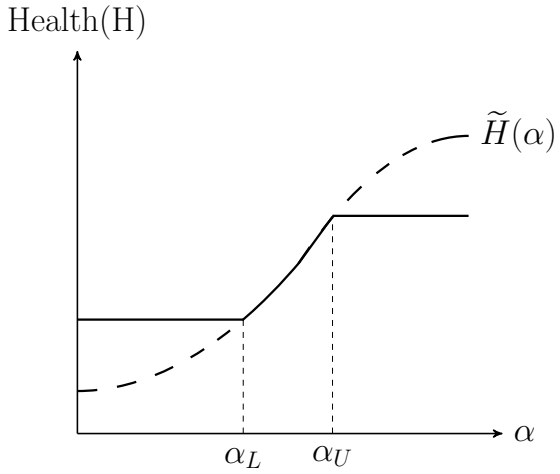
When only the binary set $\{US, MC\}$ is available, it follows directly from the weakly increasing nature of $H^*\left(\alpha\right)$ that there will be cut-off values of α such that

$$H^*\left(\alpha\right) = \begin{cases} H_L & \text{if } \alpha < \alpha_L \\ H_{US} & \text{if } \alpha \in [\alpha_L, \alpha'] \\ H_{MC} & \text{if } \alpha \in [\alpha', \alpha_U] \\ H_U & \text{if } \alpha > \alpha_U. \end{cases} \tag{B.13}$$

The introduction of female condoms expands the available technologies to the ternary set $\{US, FC, MC\}$.[29] Given that $H_{MC} > H_{FC} > H_{US}$, it follows directly that there

[29]Inserting female condoms prior to intercourse may also allow women with low bargaining

Figure B.1: Interior optimum health choices by female bargaining power



will threshold values of α such that

$$H^*(\alpha) = \begin{cases} H_L & \text{if } \alpha < \alpha_L \\ H_{US} & \text{if } \alpha \in [\alpha_L, \alpha'') \\ H_{FC} & \text{if } \alpha \in [\alpha'', \alpha'''] \\ H_{MC} & \text{if } \alpha \in [\alpha''', \alpha_U] \\ H_U & \text{if } \alpha > \alpha_U. \end{cases} \tag{B.14}$$

QED.

B.1.2 Proof of proposition two

Prior to the introduction of female condoms, the couple will only choose $s = 1$ if the set $\{US, MC\} \cap I^0$ is non-empty. Meanwhile, following the introduction of female condoms, the couple will choose $s = 1$ if the set $\{US, FC, MC\} \cap I^0$ is non-empty. Since FC is an intermediate option between US and MC, and since I^0 is a quasi-convex set, the latter condition is more likely to be satisfied. Put differently, there is a weakly positive probability that there exist couples for whom US and MC lie outside of I^0, but for

power to change the default from unprotected sex to female condom use as partners enter into bargaining over condom use.

whom $FC \in I^0$.

QED.

B.1.3 Model with transfers

We can generalize the model to include transfers in the following way. Let q_i be an action that spouse i can take, for example housework, with marginal cost to spouse i of unity and marginal benefit to the other spouse of $\phi(q_i)$. This nests the no-transfer case if $\phi(q) = 0$. Let $\phi(0) = 0$, and assume that $\phi'(q) \in [0, 1]$ and $\phi''(q) < 0$, implying that transfers involve some friction. We normalise such that at no sex, $s = 0$, both transfers are equal to zero.

The individual utility functions with sex and transfers become

$$v_i(P, H, q_i, q_{-i}) = u_i(P, H) - q_i + \phi(q_{-i}). \tag{B.15}$$

All other aspects of the model are kept intact.

Extensive Margin: The couple will choose $s = 1$ if and only if there exists some $(P, H, q_m, q_f) \in \{US, FC, MC\} \times \times R^2_+$ such that $v_i(P, H, q_i, q_{-i}) \geq u_i^0 \ \ \forall i = m, f$. It follows that the possibility of transfers increases the likelihood that $s = 1$ compared to the no-transfer case, insofar as there are cases where $s = 1$ occurs with transfers but would not if transfers were not possible. Note that it is still the case that the choice of $s = 0$ or $s = 1$ does not depend on α.

Intensive Margin: Suppose that the above condition is satisfied and thus that $s = 1$. The unconstrained household utility maximisation problem generalises to

$$\max_{H, q_m, q_f} \{(1 - \alpha)[u_m(P(H), H) - q_m + \phi(q_f)] + \alpha[u_f(P(H), H) - q_f + \phi(q_m)]\}. \tag{B.16}$$

Due to the separable form, the first-order condition with respect to H is the same for the model without transfers, namely

$$\alpha u'_{fH}\left(P(H),H\right)+\left(1-\alpha\right)u'_{mH}\left(P(H),H\right)=0. \tag{B.17}$$

Thus the unconstrained function $\widetilde{H}\left(\alpha\right)$ is preserved. In addition we now have the complementary slackness conditions

$$\left(1-\alpha\right)\geq\alpha\phi'\left(q_m\right), \tag{B.18}$$

and

$$\left(1-\alpha\right)\phi'\left(q_f\right)\leq\alpha, \tag{B.19}$$

implying a solution $\widetilde{q}_j\left(\alpha\right)$ for $j=m,f$. Note that $\phi'\left(q\right)\leq 1$) implies that only one of the complementary slackness conditions can hold with equality — i.e. q_f and q_m cannot be positive at the same time — and thus transfers will only occur in one direction. Intuitively, if α is low then $q_f>0$, and vice versa if α is high. Taken together, this gives rise to implied utilities

$$\widetilde{V}_i\left(\alpha\right)=u_i\left(P(\widetilde{H}\left(\alpha\right)),\widetilde{H}\left(\alpha\right)\right)-\widetilde{q}_i\left(\alpha\right)+\phi\left(\widetilde{q}_{-i}\left(\alpha\right)\right)i=m,f \tag{B.20}$$

with $\widetilde{V}'_f\left(\alpha\right)>0$ and $\widetilde{V}'_m\left(\alpha\right)<0$.

However, as before, if α is low enough such that $\widetilde{V}_f\left(\alpha\right)<u_f^0$ then the female's participation constraint binds. The couple instead choose an allocation that just satisfies her participation constraint, solving

$$\max_{H,q_m,q_f}\left\{U_m\left(P(H),H\right)-q_m+\phi\left(q_f\right)|U_f\left(P(H),H\right)-q_f+\phi\left(q_m\right)\geq u_f^0\right\}, \tag{B.21}$$

with the following Lagrangean

$$L=U_m\left(P(H),H\right)-q_m+\phi\left(q_f\right)+\mu_f\left\{U_f\left(P(H),H\right)-q_f+\phi\left(q_m\right)-u_f^0\right\}. \tag{B.22}$$

Since the female's participation constraint failed at the unconstrained solution, it follows that the constrained solution involves a larger implicit relative weight to the woman: $\mu_f^* \geq \alpha/(1-\alpha)$. The reverse logic applies if his participation constraint fails.

Taken together, this implies that $H^*(\alpha)$ is weakly increasing in α as in the no-transfer case, but that the range of values for which it is strictly increasing (i.e. in which an interior solution $\widetilde{H}$ is chosen) is smaller than in the no-transfer case. In terms of Figure B.1, as transfers become less costly, the horizontal segments of the line move closer to one another vertically, and thus the range $\alpha_H - \alpha_L$ becomes smaller.

B.1.4 The limiting case of frictionless transfers

Consider the limiting case where transfers are frictionless, such that $\phi'(\cdot)$ is constant and equal to unity. In this case we can simply refer to q as the net transfer from her to him, which is negative if on net he transfers to her. Hence the household's unconstrained optimisation problem collapses to

$$\max_{H,q} \{(1-\alpha)[u_m(P(H),H)+q] + \alpha[u_f(P(H),H)-q]\}. \tag{B.23}$$

It is straightforward to see that this problem has no solution, except in the knife-edge case where $\alpha = 1/2$. Taking the first-order condition with respect to q, we obtain

$$1-\alpha-\alpha = 0. \tag{B.24}$$

Since generically $\alpha \neq 1/2$, the solution will involve infinite transfers in one of the two possible directions. However, this then trivially leads to the failure of the donor's participation constraint. Suppose that $\alpha < 1/2$ whereby she is the donor. In that case the couple instead solves

$$\max_{H,q} \left\{u_m(P(H),H)+q | u_f(P(H),H)-q \geq u_f^0\right\}, \tag{B.25}$$

with Lagrangean

$$L = u_m\left(P(H), H\right) + q + \mu_f^* \left[u_f\left(P(H), H\right) - q - u_f^0\right]. \tag{B.26}$$

Note that the first-order condition with respect to q is $1 - \mu_f^* = 0$, implying $\mu_f^* = 1$. The first-order condition with respect to H therefore implies $u'_{fH}\left(P(H), H\right) = u'_{mH}\left(P(H), H\right)$. By a corresponding analysis of the case where $\alpha < 1/2$, we obtain that, with frictionless transfers, $u'_m\left(H\right) = u'_f\left(H\right)$ characterizes the couple's choice of H for any α. That is, the choice of contraceptive technology is independent of the bargaining weight. In terms of Figure B.1, we reach the limiting case where the horizontal segments of the line become completely aligned vertically, and $\widetilde{H}$ is just a constant for an value of α.

B.2 Additional Descriptive Data

Table B.1: Predictors of attrition – treatment and control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treatment		**Control**		**Test $\beta_1 = \beta_2$**			**N**	
	Mfx	**p-val**	**Mfx**	**p-val**	**χ^2**	**p-val**	**T**	**C**	**All**
Demographics									
Age in years	-0.01	0.12	-0.01	0.18	0.01	0.91	152	146	298
Years of education	-0.01	0.41	-0.01	0.52	0.03	0.87	149	146	295
Literate	-0.09	0.27	-0.09	0.32	0.02	0.88	151	144	295
Respondent is household head	-0.05	0.49	-0.01	0.90	0.19	0.66	152	146	298
Income									
Has job	-0.03	0.67	0.04	0.57	0.47	0.50	151	144	295
Personal income last 30 days (MZN)	-0.00	0.08	-0.00	0.26	1.13	0.29	152	146	298
Relationships									
In a stable relationship (incl. married)	-0.08	0.32	-0.01	0.95	0.45	0.50	152	146	298
Married (officially or unofficially)	-0.02	0.78	0.07	0.40	0.59	0.44	151	146	297
Years relation	-0.01	0.21	-0.01	0.11	0.04	0.84	121	114	235
Sexual behaviour									
HIV positive (self-report)	0.12	0.07	0.04	0.63	1.00	0.32	131	129	260
STI last 3 months (self-report)	0.06	0.47	-0.19	0.20	2.06	0.15	135	124	259
Mentions female condom as contraceptive	-0.04	0.53	-0.06	0.43	0.00	0.97	150	146	296
Baseline use									
Ever used female condom	0.05	0.60	0.12	0.32	0.07	0.80	152	146	298
Ever used male condom	0.08	0.28	-0.02	0.78	1.04	0.31	152	146	298
Ever used other	-0.07	0.27	0.03	0.69	1.15	0.28	152	146	298
Used male condom last 30 days	-0.04	0.53	-0.07	0.42	0.01	0.92	152	146	298
Current use male condom	0.07	0.23	0.02	0.82	0.55	0.46	152	146	298
Current use other	-0.03	0.68	0.08	0.26	1.09	0.30	152	146	298

Notes: N=298 in the baseline sample prior to attrition. Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Columns 1-4 show marginal effects (Mfx) and p-values (p-val) for logit regressions of the probability of attritting on each covariate, in the treatment and control group, respectively. Columns 5 and 6 show the χ^2 statistic and p-value for the test that the marginal effects are equal across the treatment and control groups. Columns 7-9 show sample sizes. Unless otherwise indicated, all are binary variables. MZN stands for Mozambican meticais. HIV stands for Human Immune-deficiency Virus. STI stands for Sexually Transmitted Infections. "Ever used other" and "Current use other" refer to use of any other modern contraceptive method apart from condoms, e.g. the pill, injectables, or an IUD.

Table B.2: Baseline balance excluding attritters

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age in years	30.80	30.65	30.93	-0.25	232	107	125
Years of education	6.30	6.36	6.25	0.27	231	107	124
Literate	0.86	0.86	0.86	-0.10	230	106	124
Respondent is household head	0.23	0.21	0.25	-0.59	232	107	125
Income							
Has job	0.37	0.41	0.34	1.10	229	105	124
Personal income last 30 days (MZN)	813.15	916.36	724.80	1.04	232	107	125
Relationships							
In a stable relationship (incl. married)	0.85	0.85	0.86	-0.12	232	107	125
Married (officially or unofficially)	0.62	0.62	0.62	-0.06	231	107	124
Years relation	9.23	9.32	9.16	0.14	184	84	100
# Partners last 12 months	0.93	0.92	0.94	-0.58	232	107	125
Sexual behaviour							
Pregnant	0.06	0.05	0.07	-0.80	231	106	125
HIV positive (self-report)	0.30	0.34	0.27	1.01	202	95	107
STI last 3 months (self-report)	0.14	0.15	0.12	0.58	205	92	113
Mentions female condoms as contraceptive	0.43	0.46	0.40	0.91	230	107	123
Contraceptive use							
Ever use female condoms	0.08	0.07	0.08	-0.15	232	107	125
Ever use male condoms	0.74	0.77	0.71	0.94	232	107	125
Ever use other	0.72	0.71	0.74	-0.43	232	107	125
Used female condoms last 30 days	0.03	0.02	0.04	-0.97	232	107	125
Used male condoms last 30 days	0.33	0.30	0.36	-0.98	232	107	125
Current use female condoms	0.03	0.03	0.03	-0.18	232	107	125
Current use male condoms	0.38	0.36	0.39	-0.43	232	107	125
Current use other	0.38	0.38	0.38	0.11	232	107	125

Notes: N=232 in the baseline sample excluding attritters (but including the 5 control respondents whose training started before endline and who are excluded from the final balanced sample, N=227). Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Column 4 presents the test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group. Unless otherwise indicated, all are binary variables. MZN stands for Mozambican meticais. HIV stands for Human Immune-deficiency Virus. STI stands for Sexually Transmitted Infections. "Ever used other" and "Current use other" refer to use of any other modern contraceptive method apart from condoms, e.g. the pill, injectables, or an IUD.

Table B.3: Baseline balance excluding attritters and 5 control respondents who started intervention before endline

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age in years	30.93	30.93	30.93	0.00	227	102	125
Years of education	6.28	6.30	6.25	0.13	227	102	125
Literate	0.86	0.85	0.86	-0.24	227	102	125
Respondent is household head	0.24	0.23	0.25	-0.40	227	102	125
Income							
Has job	0.37	0.42	0.34	1.24	227	102	125
Personal income last 30 days (MZN)	986.12	1037.75	944.00	0.32	227	102	125
Relationships							
In a stable relationship (incl. married)	0.85	0.85	0.86	-0.06	227	102	125
Married (officially or unofficially)	0.62	0.62	0.62	-0.06	227	102	125
Years relation	9.01	9.14	8.91	0.24	227	102	125
# Partners last 12 months	0.93	0.92	0.94	-0.46	227	102	125
Sexual behaviour							
Pregnant	0.06	0.04	0.07	-1.06	227	102	125
HIV positive (self-report)	0.30	0.33	0.27	0.94	197	90	107
STI last 3 months (self-report)	0.14	0.16	0.12	0.70	201	88	113
Mentions female condoms as contraceptive	0.41	0.43	0.40	0.49	227	102	125
Contraceptive use							
Ever use female condoms	0.08	0.08	0.08	-0.04	227	102	125
Ever use male condoms	0.74	0.76	0.71	0.90	227	102	125
Ever use other	0.72	0.71	0.74	-0.50	227	102	125
Used female condoms last 30 days	0.03	0.02	0.04	-0.91	227	102	125
Used male condoms last 30 days	0.33	0.29	0.36	-1.05	227	102	125
Current use female condoms	0.03	0.03	0.03	-0.11	227	102	125
Current use male condoms	0.38	0.36	0.39	-0.45	227	102	125
Current use other	0.37	0.36	0.38	-0.20	227	102	125

Notes: N=227 in the baseline sample excluding attritters and the 5 control respondents whose training started before endline and who are excluded from the final balanced sample. Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Column 4 presents the test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group. Unless otherwise indicated, all are binary variables. MZN stands for Mozambican meticais. HIV stands for Human Immune-deficiency Virus. STI stands for Sexually Transmitted Infections. "Ever used other" and "Current use other" refer to use of any other modern contraceptive method apart from condoms, e.g. the pill, injectables, or an IUD.

Table B.4: Baseline balance on covariates – diary subsample

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age in years	31.32	31.93	30.67	0.55	56	29	27
Years of education	5.95	5.45	6.48	-1.32	56	29	27
Literate	0.84	0.86	0.81	0.42	55	28	27
Respondent is household head	0.30	0.21	0.41	-1.63	56	29	27
Income							
Has job	0.38	0.38	0.37	0.07	56	29	27
Personal income last 30 days (MZN)	1005.36	927.59	1088.89	-0.39	56	29	27
Relationships							
In a stable relationship (incl. married)	0.84	0.83	0.85	-0.24	56	29	27
Married (officially or unofficially)	0.54	0.59	0.48	0.77	56	29	27
Years relation	11.78	12.82	10.58	0.86	41	22	19
# Partners last 12 months	0.91	0.86	0.96	-1.10	56	29	27
Sexual behaviour							
Pregnant	0.00	0.00	0.00	.	56	29	27
HIV positive (self-report)	0.33	0.38	0.29	0.60	48	24	24
STI last 3 months (self-report)	0.12	0.12	0.12	0.00	48	24	24
Mentions female condoms as contraceptive	0.27	0.31	0.23	0.66	55	29	26

Notes: N=56 diaries respondents in the balanced panel. Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Column 4 presents the test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group. Unless otherwise indicated, all are binary variables. MZN stands for Mozambican meticais. HIV stands for Human Immune-deficiency Virus. STI stands for Sexually Transmitted Infections.

Table B.5: Baseline balance on use – diary subsample

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Ever use (survey)							
Ever used female condoms	0.05	0.07	0.04	0.53	56	29	27
Ever used male condoms	0.77	0.79	0.74	0.45	56	29	27
Ever used other	0.84	0.86	0.81	0.47	56	29	27
Use last 30 days (survey)							
Used female condoms last 30 days	0.04	0.00	0.07	-1.44	56	29	27
Used male condoms last 30 days	0.36	0.28	0.44	-1.31	56	29	27
Current use (survey)							
Current use female condoms	0.02	0.00	0.04	-1.00	56	29	27
Current use male condoms	0.39	0.38	0.41	-0.21	56	29	27
Current use other	0.48	0.41	0.56	-1.05	56	29	27
Ever used (diaries)							
Used female condoms in baseline weeks	0.04	0.03	0.04	-0.05	56	29	27
Used male condoms in baseline weeks	0.45	0.41	0.48	-0.50	56	29	27
Use last 30 days (diaries)							
Used female condoms in baseline last 30 days	0.04	0.03	0.04	-0.05	56	29	27
Used male condoms in baseline last 30 days	0.54	0.52	0.56	-0.28	56	29	27
Use last 14 days (diaries)							
Used female condoms in baseline last 14 days	0.04	0.03	0.04	-0.05	56	29	27
Used male condoms in baseline last 14 days	0.39	0.34	0.44	-0.75	56	29	27
Discussions and sex acts (diaries)							
% sex acts with discussion about condom use	0.16	0.13	0.19	-0.64	56	29	27
% sex acts with female-initiated discussion about condom use	0.13	0.07	0.19	-1.55	56	29	27
# sex acts per week	0.90	0.89	0.91	-0.04	55	28	27

Notes: N=56 diaries respondents in the balanced panel. Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Column 4 presents the test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group. "Ever used other" and "Current use other" refer to use of any other modern contraceptive method apart from condoms, e.g. the pill, injectables, or an IUD. "% sex acts" represents the percentage of sex acts per individual during which partners discussed condom use. "# sex acts per week" represents the average number of sex acts per week per individual.

Table B.6: Diary sample representativeness of full sample – covariates

	Mean	Survey Mean	Diaries Mean	t-test	Survey N	Diaries N
Demographics						
Age in years	30.48	30.32	31.32	-0.81	298	56
Years of education	6.17	6.21	5.95	0.62	295	56
Literate	0.84	0.84	0.84	0.14	295	55
Respondent is household head	0.24	0.22	0.30	-1.18	298	56
Income						
Has job	0.38	0.38	0.38	0.02	295	56
Personal income last 30 days (MZN)	786.90	745.85	1005.36	-1.18	298	56
Relationships						
In a stable relationship (incl. married)	0.84	0.85	0.84	0.12	298	56
Married (officially or unofficially)	0.61	0.63	0.54	1.24	297	56
Years relation	9.13	8.66	11.78	-2.24**	235	41
# Partners last 12 months	0.92	0.92	0.91	0.24	298	56
Sexual behaviour						
Pregnant	0.05	0.05	0.00	4.11***	297	56
HIV positive (self-report)	0.33	0.33	0.33	-0.09	260	48
STI last 3 months (self-report)	0.13	0.13	0.12	0.12	259	48
Mentions female condom as contraceptive	0.39	0.41	0.27	2.08**	296	55
Bargaining power (principle components)						
Assets score 1	0.76	0.76	0.75	0.08	263	49
Assets score 2	0.43	0.43	0.45	-0.15	263	49
Assets score 3	1.17	1.18	1.14	0.40	263	49
Decision-making	1.82	1.79	2.01	-1.43	235	46
Power dynamics	2.84	2.84	2.84	-0.04	235	46

Notes: N=298 in the baseline sample, of which N=56 are in the subsample who respond to the diaries. Lower sample sizes in columns 5 and 6 reflect observations that are missing or not applicable. "Survey" contains all individuals in the balanced panel, whether or not they participated in the diaries. "Diaries" contains only the subsample of individuals who also responded to the diaries. Column 4 presents the t-test statistic for the null hypothesis that the mean in the diary subsample is equal to the mean in the survey sample. Unless otherwise indicated, all are binary variables. MZN stands for Mozambican meticais. HIV stands for Human Immune-deficiency Virus. STI stands for Sexually Transmitted Infections. "Assets score 1", "Assets score 2" and "Assets score 3" represent the three principal components loaded by the assets module as identified in Table B.8. "Decision-making" and "Power dynamics" represent the two principal components loaded by each of these survey modules as identified in Table B.9. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.7: Diary sample representativeness of full sample – baseline use

	Mean	Survey Mean	Diaries Mean	t-test	Survey N	Diaries N
Ever use (baseline survey)						
Ever used female condoms	0.08	0.09	0.05	0.98	298	56
Ever used male condoms	0.75	0.74	0.77	-0.37	298	56
Ever used other	0.74	0.72	0.84	-2.16**	298	56
Use last 30 days (baseline survey)						
Used female condoms last 30 days	0.03	0.03	0.04	-0.33	298	56
Used male condoms last 30 days	0.32	0.32	0.36	-0.60	298	56
Current use (baseline survey)						
Current use female condoms	0.02	0.02	0.02	0.27	298	56
Current use male condoms	0.39	0.39	0.39	-0.090	298	56
Current use other	0.40	0.39	0.48	-1.27	298	56

Notes: N=298 in the baseline sample, of which N=56 are in the subsample who respond to the diaries. "Survey" contains all individuals in the balanced panel, whether or not they participated in the diaries. "Diaries" contains only the subsample of individuals who also responded to the diaries. Column 4 presents the t-test statistic for the null hypothesis that the mean in the diary subsample is equal to the mean in the survey sample. "Ever used other" and "Current use other" refer to use of any other modern contraceptive method apart from condoms, e.g. the pill, injectables, or an IUD. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.8: Assets brought to the relationship – principal component analysis

	(PC1)	(PC2)	(PC3)	Unexplained
Assets brought to the relationship				
Jewellery			0.4091	0.2922
Livestock			0.7958	0.2609
Land		0.5964		0.4774
Electronic appliances		0.7373		0.3262
Money	0.6782			0.2543
Mobile phone	0.6587			0.2568
Kitchenware			0.4131	0.3372

Notes: All values taken from the baseline survey (N=298), restricted to all women who lived with their partner at baseline, including a few who reported not be in a *stable* relationship with this partner (N=264). Factor loadings from a Principal Component Analysis of all questions in the assets module simultaneously (i.e. "Did you bring any of the following assets to your relationship?"), as described in Table 3. Only loadings greater than or equal to 0.25 are displayed. All variables are dummy variables indicating whether specific asset was brought to relationship, such that a higher value proxies greater bargaining power for the respondent. As shown, the asset questions load three separate components (PC1, PC2, PC3), all of which are used in the bargaining power heterogeneity analysis.

Table B.9: Bargaining power – principal component analysis

	(PC1)	(PC2)	Unexplained
Decision-making			
Clothes	0.2728		0.6316
Phone credit	0.2678		0.6491
Children's education	0.3583		0.3933
Her health	0.3468		0.4190
Children's health	0.3288		0.4757
Her employment	0.3390		0.4288
Spending earnings	0.3216		0.5115
Visiting friends	0.3586		0.3648
Visiting family	0.3436		0.4152
In general, more say			0.8677
In general, more power			0.9097
Power dynamics			
We do what he wants		0.3039	0.7457
He won't let me wear certain things			0.8812
I'm quiet around him		0.3645	0.6395
He has more say about joint decisions		0.3081	0.7497
He controls who I spend time with		0.2915	0.7666
I feel trapped or stuck		0.2926	0.7293
He does what he wants		0.3173	0.7402
I'm more committed		0.2984	0.7438
He sees other people			0.8898
He wants to know where I am		0.2775	0.8013
He gets his way when we disagree			0.9082

Notes: Factor loadings from a Principal Component Analysis of all questions in the "decision-making" and "power dynamics" modules simultaneously, as described in Table 3. The decision-making module was enumerated to all respondents (N=298), except the questions "who has more say" and "who has more power" which were asked only of women in a stable relationship at baseline (N=250). Decision-making variables are indicators for whether respondent was involved in making decisions on each of the activities or if respondent had more say/more power than her partner."Power dynamics" questions were only asked from women who were in a stable relationship at baseline (N=250). Likert-scale answers range from 1 (completely disagree) to 4 (completely agree), and are recoded such that a greater value represents higher bargaining power for the respondent. Only loadings greater than or equal to 0.25 are displayed. As shown, the two modules load two separate components (PC1, PC2), which are both used in the bargaining power heterogeneity analysis.

Table B.10: Balance – principal components of bargaining power

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Assets 1	0.76	0.81	0.71	0.82	263	128	135
Assets 2	0.43	0.41	0.45	-0.33	263	128	135
Assets 3	1.18	1.24	1.13	0.94	263	128	135
Decision-making	1.79	1.83	1.75	0.60	235	114	121
Power dynamics	2.84	2.82	2.85	-0.17	235	114	121

Notes: All values taken from the baseline survey (N=298). "Treatment" contains all individuals assigned to the treatment group (i.e. to the first round of the family planning training sessions), whether or not they attended the sessions. "Control" contains all individuals assigned to the control group (i.e. to the second round of training sessions). Column 4 presents the test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group. "Assets score 1", "Assets score 2" and "Assets score 3" represent the three principal components loaded by the assets module as identified in Table B.8. It was enumerated to all women who lived with their partner at baseline, including a few who did not claim to be in a stable relationship (N=264). "Decision-making" and "Power dynamics" represent the two principal components loaded by each of these survey modules as identified in Table B.9. The decision-making module was enumerated to all respondents (N=298), except the questions "who has more say" and "who has more power" which were asked only of women in a stable relationship at baseline (N=250). "Power dynamics" questions were also only asked to women who were in a stable relationship at baseline (N=250). Any lower sample sizes reflect values missing or unwillingness to answer. All variables are (re-)coded such that a higher value proxies greater bargaining power for the respondent. All components are scaled such that the least empowered woman on that component has a score of zero. They are also normalised such that a one point increase in each component represents an increase of one standard deviation.

B.3 Additional Analyses

Table B.11: Treatment effects – primary outcome variables – OLS specification

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Use last 30 days female condoms**	(5) **Use last 30 days male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	0.190***	-0.027	0.035	0.046**	-0.028	0.078**	0.071	0.033
	(0.046)	(0.050)	(0.056)	(0.023)	(0.059)	(0.030)	(0.060)	(0.063)
Observations	227	227	227	227	227	227	227	227
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Regressions on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill, injectables or IUD. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days (this was only asked for condoms, not for other contraceptive methods), and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model specifications. The specifications are a replication of Table A.1 without including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Randomisation inference p-values [r.i. p-value] are estimated from Monte Carlo simulations re-assigning treatment within facilitator strata, with 1000 repetitions. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.12: Treatment effects on female condom use, by baseline male condom use – OLS specification

	(1) **Ever use female condom** **No use male condom at baseline**	(2) **Ever use female condom** **Current use male condom at baseline**	(3) **Last 30 days female condom** **No use male condom at baseline**	(4) **Last 30 days female condom** **Current use male condom at baseline**	(5) **Current use female condom** **No use male condom at baseline**	(6) **Current use female condom** **Current use male condom at baseline**
Treatment	0.151***	0.264***	0.072**	0.026	0.092***	0.056
	(0.052)	(0.081)	(0.030)	(0.034)	(0.033)	(0.056)
Observations	141	86	141	86	141	86
Control mean endline	0.092	0.081	0.000	0.027	0.000	0.054

Notes: Regressions on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms: ever used in columns 1-2, used in last 30 days in columns 3-4, and currently using in columns 5-6. Odd-numbered columns present results for the subsample of individuals who were not currently using male condoms (No use) at baseline; even-numbered columns present results for the subsample of individuals who were currently using male condoms (Current use) at baseline. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model specifications. The specifications are a replication of Table A.2 without including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation. Randomisation inference p-values [r.i. p-value] are estimated from Monte Carlo simulations re-assigning treatment within facilitator strata, with 1000 repetitions. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.13: Impacts on current use of female condoms – heterogeneity by bargaining power – OLS specification

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.129***	0.112***	0.160***	0.229**	0.205*
	(0.045)	(0.039)	(0.051)	(0.095)	(0.116)
Assets score 1	-0.002				
	(0.009)				
Treatment*Assets score 1	-0.044**				
	(0.017)				
Assets 2		0.003			
		(0.009)			
Treatment*Assets score 2		-0.031**			
		(0.015)			
Assets 3			-0.002		
			(0.005)		
Treatment*Assets score 3			-0.055***		
			(0.019)		
Decision-making				-0.013	
				(0.021)	
Treatment*Decision-making				-0.081**	
				(0.039)	
Power dynamics					0.016
					(0.018)
Treatment*Power dynamics					-0.041
					(0.040)
Controls	✓	✓	✓	✓	✓
Observations	201	201	201	182	182
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Based on balanced sample of respondents (N=227) with stable relationship at baseline, N=206; N=201 of them have non-missing values on all assets and control variables; N=182 of them have non-missing values on all decision-making, power dynamics, and control variables. Dependent variables are binary indicators for current use of female condoms (FC). "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as described in Table 3 and identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as described in Table 3 and identified in Table B.9. For comparability the components are scaled so that the woman with least bargaining power on that measure has a score of zero. The components are normalized such that a one point increase in each measure represents an increase of one standard deviation. All regressions are linear probability model specifications. The specifications are a replication of Table A.4 without including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.14: Treatment effects on male condom use, by baseline male condom use

	(1) **Ever use male condom** **No use male condom at baseline**	(2) **Ever use male condom** **Current use male condom at baseline**	(3) **Last 30 days male condom** **No use male condom at baseline**	(4) **Last 30 days male condom** **Current use male condom at baseline**	(5) **Current use male condom** **No use male condom at baseline**	(6) **Current use male condom** **Current use male condom at baseline**
Treatment	-0.000	0.002	0.032	-0.158	0.101	-0.002
	(0.059)	(0.053)	(0.065)	(0.097)	(0.068)	(0.105)
Facilitator indicators	✓	✓	✓	✓	✓	✓
Observations	141	86	141	86	141	86
Control mean endline	0.754	0.946	0.231	0.595	0.215	0.595
$\chi^2(1) : (a) = (b)$		0.00		2.65		0.68
$\Pr > \chi^2$		0.973		0.103		0.411

Notes: Regressions on the balanced sample, N=227. Dependent variables are binary indicators for the use of male condoms (MC): ever used in columns 1-2, used in last 30 days in columns 3-4, and currently using in columns 5-6. Odd-numbered columns present results for the subsample of individuals who were not currently using male condoms (No use) at baseline; even-numbered columns present results for the subsample of individuals who were currently using male condoms (Current use) at baseline. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.15: Treatment effects – other outcome variables

	(1) **HIV knowledge (score 0-6)**	(2) **HIV positive**	(3) **STI last 3 months**	(4) **Well-being (score 0-12)**	(5) **Violence (score 0-5)**
Treatment	-0.136*	-0.021	0.021	0.171	0.076
	(0.081)	(0.042)	(0.033)	(0.260)	(0.195)
Facilitator f.e.'s	✓	✓	✓	✓	✓
Observations	219	196	185	212	162
Control mean endline	5.758	0.313	0.054	8.135	1.149

Notes: Regressions on the balanced sample (N=227), except for the violence outcome since these questions were only enumerated to women in a stable relationship at baseline. Missing observations reflect not applicable, does not want to answer, and cases where the facilitator indicator perfectly predicts the outcome variable. Dependent variables are as follows, all measured at endline: column 1, a score from six questions testing knowledge about how HIV can and cannot be transmitted; column 2, a self-reported dummy for HIV-positive status; column 3, a self-reported dummy for having had an STI in the last three months; column 4, a score from twelve questions on well-being (higher scores indicate greater well-being); column 5 a score from five questions about emotional and physical violence (a higher score indicates greater violence). "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.16: Treatment effects – heterogeneity by relationship status

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Use last 30 days female condoms**	(4) **Use last 30 days male condoms**	(5) **Current use female condoms**	(6) **Current use male condoms**
Treatment	0.358***	-0.089	0.040	0.061	0.165*	0.179
	(0.103)	(0.112)	(0.054)	(0.153)	(0.088)	(0.150)
Stable relationship	0.030	-0.038	0.007	-0.052	0.024	-0.064
	(0.051)	(0.078)	(0.020)	(0.120)	(0.024)	(0.109)
Treat×Stable relationship	-0.202*	0.090	0.009	-0.132	-0.102	-0.141
	(0.109)	(0.121)	(0.064)	(0.166)	(0.093)	(0.162)
Facilitator f.e.'s	✓	✓	✓	✓	✓	✓
Observations	227	227	220	221	227	227
Control mean endline	0.088	0.824	0.010	0.366	0.020	0.353

Notes: Regressions on the balanced sample, N=227. Reduced observations in columns (3) and (4) reflect there being no variation in the outcome variable conditional on the facilitator fixed effect and controls. Dependent variables are binary indicators for the use of female condoms (FC) and male condoms (MC). Columns 1 and 2 refer to whether the respondent has ever used the method, columns 3 and 4 to whether she has used it in the last 30 days, and columns 5 and 6 to whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Stable relationship" is a dummy equal to one if the respondent reports being in a stable relationship at baseline. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.17: Impacts on current use of female condoms – LASSO selects heterogeneity by bargaining power

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.446***	0.384**	0.427***	0.550***	0.506
	(0.161)	(0.156)	(0.160)	(0.181)	(0.199)
Assets score 1	-0.009				
	(0.007)				
Treatment*Assets score 1	-0.038**				
	(0.016)				
Assets score 2					
Treatment*Assets score 2					
Assets score 3					
Treatment*Assets score 3			-0.056***		
			(0.020)		
Decision-making					
Treatment*Decision-making				-0.104***	
				(0.036)	
Power dynamics					0.019
					(0.018)
Treatment*Power dynamics					-0.042
					(0.041)
Age	0.001	0.001	0.001	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Years of education	0.003	0.002	0.002	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
Has job				-0.033	
				(0.028)	
Treatment*Age	-0.008**	-0.007**	-0.007**	-0.006*	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Treatment*Years of education	-0.011	-0.012	-0.008	-0.011	-0.007
	(0.011)		(0.011)	(0.009)	(0.010)
Treatment*Literacy	0.052	0.058	0.055	0.032	0.019
	(0.111)	(0.114)	(0.111)	(0.112)	(0.128)
Treatment*high HIV risk perception	- 0.142*	-0.135*	-0.141*	-0.134*	-0.107
	(0.080)	(0.100)	(0.080)	(0.080)	(0.079)
Treatment*Walking distance to health centre	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	201	201	201	182	182
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Estimations with post-double LASSO based on balanced sample (N=227) with stable relationship at baseline, N=206; N=201 of them have non-missing values on all assets and control variables; N=182 of them have non-missing values all decision-making, power dynamics, and control variables. Dependent variables are binary indicators for current use of female condoms at endline. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect."Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. In each of the columns we run the post-double LASSO with one of the five principal components, its interaction with treatment, and all controls and their interactions with treatment. The coefficients presented in that column are the ones that were selected by the post-double LASSO. The Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.18: Treatment effects – bargaining power

	Mfx	s.e.	p-val	N
Who decides about...				
...buying clothes for you?	-0.03	0.04	0.46	227
...buying phone credit?	0.03	0.04	0.52	227
...education for the children?	-0.03	0.04	0.46	226
...health expenses for you?	-0.10	0.04	0.01	227
...health expenses for the children?	-0.06	0.04	0.13	225
...if you are allowed to work?	-0.06	0.04	0.16	227
...how earnings are used?	-0.01	0.04	0.74	227
...visits to friends?	-0.00	0.04	1.00	226
...visits to family?	-0.01	0.05	0.80	226
Who usually has more say when you talk about serious things	0.11	0.05	0.03	177
In general, who do you think has more power in your relationship	0.11	0.05	0.02	177
Power dynamics				
Most of the time, we do what my partner wants to do	-0.03	0.05	0.45	193
My partner won't let me wear certain things	-0.01	0.05	0.82	193
When my partner and I are together, I'm pretty quiet	-0.04	0.05	0.37	193
My partner has more say about important decisions that affect us	-0.03	0.05	0.51	193
My partner tells me who I can spend time with	-0.03	0.05	0.52	193
I feel trapped or stuck in our relationship	-0.00	0.05	0.99	193
My partner does what he wants, even if I do not want him to	-0.05	0.05	0.27	193
I am more committed to our relationship than my partner is	0.04	0.05	0.34	193
My partner is involved with other people apart from me	-0.15	0.05	0.00	193
My partner always wants to know where I am	0.13	0.04	0.00	193
When my partner and I disagree, he gets his way most of the time	0.07	0.05	0.12	193

Notes: Regressions on the balanced sample (N=227). Lower sample sizes reflect observations that are missing or not applicable. Dependent variables are the individual bargaining power indicators measured at endline, as indicated in each row. The decision-making questions "Who has more say" and "Who has more power" as well as the "Power dynamics" questions were asked only of women in a stable relationship. Columns 1 and 2 shows the marginal effects (Mfx), standard errors (s.e.), and p-values (p-val) respectively, for regressions on the "Treatment" indicator of being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17), since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table B.19: Current use of female condoms at endline – correlation with bargaining power among women using male condoms at baseline

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Assets 1	-0.047**				
	(0.019)				
Assets 2		-0.021**			
		(0.010)			
Assets 3			-0.060**		
			(0.027)		
Decision-making				-0.117***	
				(0.044)	
Power dynamics					-0.040
					(0.043)
Controls	✓	✓	✓	✓	✓
Observations	75	75	75	75	75
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions based on balanced sample (N=227), restricted to women who reported current use of male condoms at baseline, and who answer the assets module (columns 1-3) or all two survey modules on bargaining power (columns 4-5) (N=75). Dependent variables are binary indicators for whether the respondent is currently using female condoms at endline. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and"Power dynamics" are the first two principal components from these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. We omit the interaction terms between the measures of bargaining power and treatment, because power is severely weakened by the loss in sample size when we restrict attention to only those women who were using male condoms at baseline. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.20: Impacts on current use of male condoms – heterogeneity by bargaining power

	(1) Current use male condoms	(2) Current use male condoms	(3) Current use male condoms	(4) Current use male condoms	(5) Current use male condoms
Treatment	-0.006	0.036	-0.026	0.026	0.093
	(0.076)	(0.065)	(0.106)	(0.142)	(0.183)
Assets score 1	-0.067**				
	(0.032)				
Treatment*Assets score 1	0.098*				
	(0.055)				
Assets score 2		-0.050			
		(0.032)			
Treatment*Assets score 2		0.080*			
		(0.041)			
Assets score 3			-0.046		
			(0.029)		
Treatment*Assets score 3			0.082		
			(0.078)		
Decision-making				0.015	
				(0.050)	
Treatment*Decision-making				0.008	
				(0.064)	
Power dynamics					0.035
					(0.043)
Treatment*Power dynamics					-0.019
					(0.061)
Controls	✓	✓	✓	✓	✓
Observations	201	201	201	182	182
Control mean endline	0.353	0.353	0.353	0.353	0.353

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=201) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182). Dependent variables are binary indicators for current use of male condoms at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.21: Treatment effects – heterogeneity by baseline bargaining power, controlling for baseline use of other contraceptives and its interaction with treatment

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.124**	0.110**	0.154***	0.207**	0.179
	(0.050)	(0.047)	(0.055)	(0.098)	(0.109)
Current use other (baseline)	-0.001	-0.003	-0.003	-0.029	-0.035
	(0.037)	(0.038)	(0.036)	(0.026)	(0.029)
Treatment*Current use other (baseline)	0.015	0.007	0.018	0.039	0.060
	(0.077)	(0.078)	(0.077)	(0.074)	(0.078)
Treatment*Assets score 1	-0.045**				
	(0.018)				
Assets score 1					
Treatment*Assets score 2		-0.032**			
		(0.016)			
Assets score 2					
Treatment*Assets score 3			-0.056***		
			(0.019)		
Assets score 3					
Treatment*Decision-making				-0.077**	
				(0.038)	
Decision-making					
Treatment*Power dynamics					-0.040
					(0.040)
Power dynamics					
Controls	✓	✓	✓	✓	✓
Observations	201	201	201	182	182
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=201) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182). Dependent variables are binary indicators for current use of female condoms at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Current use other (baseline)" is a dummy equal to one if the respondent reports using any non-condom forms of modern contraception at baseline, mainly the pill or injectables. "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.22: Treatment effects – heterogeneity by baseline bargaining power, controlling for HIV status and its interaction with treatment

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.094*	0.080*	0.124**	0.225**	0.211
	(0.051)	(0.043)	(0.059)	(0.089)	(0.137)
HIV positive (self-report)	0.021	0.023	0.024	0.047	0.047
	(0.042)	(0.041)	(0.042)	(0.045)	(0.046)
Treatment*HIV positive	0.081	0.083	0.079	0.067	0.035
	(0.098)	(0.098)	(0.099)	(0.101)	(0.103)
Treatment*Assets score 1	-0.035*				
	(0.019)				
Assets score 1					
Treatment*Assets score 2		-0.025			
		(0.015)			
Assets score 2					
Treatment*Assets score 3			-0.048**		
			(0.019)		
Assets score 3					
Treatment*Decision-making				-0.086**	
				(0.042)	
Decision-making					
Treatment*Power dynamics					-0.043
					(0.046)
Power dynamics					
Controls	✓	✓	✓	✓	✓
Observations	173	173	173	158	158
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=201) as well as on self-reported HIV status (N=173) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182) as well as on self-reported HIV status (N=158) . Dependent variables are binary indicators for current use of female condoms (FC) at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "HIV positive (self-report)" is a dummy equal to one if the respondent reports being HIV-positive at baseline. "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.23: Treatment effects – heterogeneity by baseline bargaining power, controlling for risk beliefs and its interaction with treatment

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.250***	0.224***	0.281***	0.340***	0.274**
	(0.086)	(0.081)	(0.090)	(0.120)	(0.126)
Believes high risk to self	-0.030	-0.030	-0.035	-0.002	-0.007
	(0.042)	(0.043)	(0.042)	(0.035)	(0.038)
Treatment*Believes high risk to self	-0.150*	-0.143*	-0.146*	-0.134*	-0.098
	(0.081)	(0.081)	(0.081)	(0.075)	(0.079)
Treatment*Assets score 1	-0.051***				
	(0.020)				
Assets score 1	0.001				
	(0.011)				
Treatment*Assets score 2		-0.034**			
		(0.016)			
Assets score 2		0.008			
		(0.011)			
Treatment*Assets score 3			-0.061***		
			(0.021)		
Assets score 3			0.001		
			(0.005)		
Treatment*Decision-making				-0.089**	
				(0.038)	
Decision-making				-0.017	
				(0.021)	
Treatment*Power dynamics					-0.039
					(0.041)
Power dynamics					0.015
					(0.020)
Controls	✓	✓	✓	✓	✓
Observations	200	200	200	181	181
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=206) as well as on beliefs about HIV risk for self (N=200) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182) as well as on beliefs about HIV risk for self (N=181) . Dependent variables are binary indicators for current use of female condoms at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Believes high risk to self" is an indicator equal to one if the respondent's answer to the question "what is your risk of contracting HIV/AIDS in a case of unprotected sex?" was above the median on a 1-5 scale. In practice this corresponds to an answer of "5, very risky" since the median response was "4, risky". "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.24: Treatment effects – heterogeneity by baseline bargaining power, controlling for preferences to work and its interaction with treatment

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.132**	0.109**	0.155***	0.250**	0.211*
	(0.052)	(0.044)	(0.057)	(0.102)	(0.119)
Prefers to work	0.016	0.010	0.013	0.048*	0.061
	(0.036)	(0.037)	(0.038)	(0.027)	(0.041)
Treatment*Prefers to work	-0.010	0.013	0.006	0.048	-0.008
	(0.104)	(0.108)	(0.106)	(0.108)	(0.116)
Assets score 1	-0.003				
	(0.009)				
Treatment*Assets score 1	-0.053**				
	(0.022)				
Assets score 2		0.003			
		(0.009)			
Treatment*Assets score 2		-0.029**			
		(0.014)			
Assets score 3			-0.002		
			(0.006)		
Treatment*Assets score 3			-0.053**		
			(0.021)		
Decision-making				-0.029	
				(0.022)	
Treatment*Decision-making				-0.092**	
				(0.041)	
Power dynamics					0.018
					(0.023)
Treatment*Power dynamics					-0.040
					(0.045)
Controls	✓	✓	✓	✓	✓
Observations	183	183	183	163	163
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=206) as well as on beliefs about HIV risk for self (N=200) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182) as well as on beliefs about HIV risk for self (N=181) . Dependent variables are binary indicators for current use of female condoms at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Believes high risk to self" is an indicator equal to one if the respondent's answer to the question "what is your risk of contracting HIV/AIDS in a case of unprotected sex?" was above the median on a 1-5 scale. In practice this corresponds to an answer of "5, very risky" since the median response was "4, risky". "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.25: Treatment effects – heterogeneity by baseline bargaining power, controlling for beliefs about their partner being involved with other women and its interaction with treatment

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	0.122**	0.106**	0.148**	0.246**	0.224*
	(0.055)	(0.051)	(0.073)	(0.103)	(0.129)
Beliefs partner is involved	-0.005	-0.012	-0.005	0.012	0.011
	(0.050)	(0.051)	(0.051)	(0.046)	(0.051)
Treatment*Beliefs partner is involved	-0.001	-0.006	-0.014	-0.029	0.001
	(0.079)	(0.080)	(0.081)	(0.074)	(0.080)
Assets score 1	-0.008				
	(0.010)				
Treatment*Assets score 1	-0.040**				
	(0.018)				
Assets score 2		-0.010			
		(0.011)			
Treatment*Assets score 2		-0.020			
		(0.015)			
Assets score 3			-0.015		
			(0.017)		
Treatment*Assets score 3			-0.043*		
			(0.025)		
Decision-making				-0.021	
				(0.022)	
Treatment*score Decision-making				-0.085**	
				(0.041)	
Power dynamics					0.020
					(0.022)
Treatment*Power dynamics					-0.045
					(0.043)
Controls	✓	✓	✓	✓	✓
Observations	167	167	167	169	169
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Regressions on balanced sample (N=227), restricted to women with stable relationship at baseline who have non-missing values on all assets and control variables (N=206) as well as on beliefs about their partner being involved with other women (N=167) or who have non-missing values on all decision-making, power dynamics, and control variables (N=182) as well as on beliefs about their partner being involved with other women (N=16) . Dependent variables are binary indicators for current use of female condoms at endline. "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Believes partner is involved" is an indicator equal to one if the respondent's answer to the question "My partner is involved with other people apart from me?" was "Totally agree" or "Agree" and 0 if the answer was "Disagree" or "Totally disagree". "Assets score 1", "Assets score 2" and "Assets score 3" are the first three principal components from the assets module, as identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors (s.e.) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, years of education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.26: Treatment effects – primary outcome variables – inverse probability weighting

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	0.196***	-0.014	0.028	0.046**	-0.052	0.081***	0.044	0.055
	(0.043)	(0.041)	(0.044)	(0.022)	(0.057)	(0.031)	(0.059)	(0.054)
Observations	227	227	227	227	227	227	227	227
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Regressions on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill or injectables. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days, and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Observations are weighted by the inverse probability of them appearing in the endline sample, as predicted from a logit estimation regressing non-attrition on treatment status and all variables in Table 1, except self-reported HIV and STI status which have missing values that are likely to be non-random. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.27: Lee bounds – primary outcome variables

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Upper	0.223***	0.090	0.169*	0.066**	0.005	0.112***	0.137*	0.106
	(0.058)	(0.084)	(0.086)	(0.028)	(0.073)	(0.037)	(0.077)	(0.079)
Lower	0.045	-0.087	-0.008	-0.010	-0.173**	-0.020	-0.040	-0.071
	(0.079)	(0.061)	(0.065)	(0.010)	(0.087)	(0.014)	(0.084)	(0.084)
95% C.I. Upper bound	0.318	0.228	0.311	0.112	0.124	0.173	0.264	0.237
95% C.I. Lower bound	-0.084	-0.187	-0.115	-0.026	-0.315	-0.042	-0.178	-0.209
Proportion trimmed	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150
Observations	298	298	298	298	298	298	298	298
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Regressions on the baseline sample, N=298. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill or injectables. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days, and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Bounds do not include facilitator indicators, as attrition is not monotonic on treatment status conditional on facilitator indicators. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

B.4 Diary analyses

As robustness checks, we also re-estimate the individual-level impacts using the diary data. First, to check that the diary respondents are representative of the full sample in terms of the impacts estimated from the survey data, we repeat estimation of Equation 3 with the survey data but only for diary respondents. Table B.28 confirms that the results are not different from the main sample, although the result for female condom use in the last 30 days loses significance due to the loss of sample size.

Next, we re-estimate impacts using the main outcome variables measured from the diary data. We first use the linear probability fixed effects panel model from Equation B.29. Table B.29 show the results.[30] Again, we see significant impacts for the use of female condoms during the full endline period and the last 30 days, while we see no significant impacts on male condom use and on female condom use in the last 14 days.

We also estimate ANCOVA specifications, for comparability with the main results (Table B.30).[31] Consistent with the survey data, we see significant and positive impacts on ever use and use in the last 30 days of female condoms. The impact on use of female condoms in the last 14 days is no longer significant due to the loss of sample size. We see no significant impacts for the use of male condoms.

[30] Again, similar results are obtained with a logit specification (see B.44).

[31] Since the nature of the diary data is such that baseline observations are missing for some respondents, we follow McKenzie (2015) and estimate:

$$Pr\left[Y_{if1}=1|Y_{if0},treat_{if},\eta_f\right]=\alpha+\delta Y_{if0}+\beta_1 treat_{if}+\gamma\text{missbase}_{if}+\sigma\text{missbase}_{if}\times Y_{if0}+\eta_f \tag{B.27}$$

where Y_{if1} is the outcome variable of interest for the endline period. Y_{if0} is the value of the outcome variable for the baseline period. $missbase_{if}$ is a dummy equal to one if the respondent is missing the value of the outcome variable during the baseline period. $missbase_{if} \times Y_{if0}$ then sets the baseline value to zero in the case that it is missing. Inclusion of this dummy means that δ is estimated only for respondents whose baseline data is not missing. $treat_{if}$ is again a dummy for being in the treatment group, and η_f is again a facilitator fixed effect. β_1 represents the intent-to-treat effect, this time as estimated on the diary subsample.

Table B.28: Treatment effects – survey variables, diary subsample

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	0.249***	-0.040	0.040	0.087	-0.107	0.125*	0.069	0.068
	(0.094)	(0.093)	(0.066)	(0.059)	(0.132)	(0.068)	(0.133)	(0.108)
Observations	50	50	50	50	50	50	50	50
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Regressions on the balanced sample of diary respondents who have answered the endline survey (N=50). Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill, injectables or IUD. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days, and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Facilitator indiciators are dropped because facilitator perfectly predicts outcomes for many observations in this subsample. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.29: Impacts on proportion of sex acts in a week that are protected by male and female condoms – diary subsample

	(1) **female condoms full endline**	(2) **male condoms full endline**	(3) **female condoms last 30 days**	(4) **male condoms last 30 days**	(5) **female condoms last 14 days**	(6) **male condoms last 14 days**
Treat×endline	0.120**	-0.148	0.123**	-0.033	0.054	-0.099
	(0.054)	(0.093)	(0.050)	(0.103)	(0.045)	(0.113)
Facilitator× endline f.e.'s	✓	✓	✓	✓	✓	✓
Observations	383	383	252	252	175	175
Control mean	0.020	0.350	0.015	0.374	0.021	0.412

Notes: Regressions on the balanced diary sample, N=56. Dependent variables are the proportion of sex acts of a respondent in a particular week that are protected by male and female condoms. Column 1 and 2 report the results for the full endline period, Column 3 and 4 for the last 30 days, and Column 5 and 6 the last 14 days. All regressions are linear probability individual fixed effects models with the respondent-week as the unit of observation, and weeks with zero sex acts are counted as missing (N=383 for the full endline period, N=252 for the last 30 days, and N=175 for the last 14 days). "Treat × endline" is an indicator for observations in the treatment group (i.e. to the first round of the family planning training sessions) during the relevant endline period ("full endline", "last 30 days", or "last 14 days") as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions include facilitator × endline fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$

Table B.30: Impacts on condom use – ANCOVA specification, diary subsample

	(1) **female condoms full endline**	(2) **male condoms full endline**	(3) **female condoms last 30 days**	(4) **male condoms last 30 days**	(5) **female condoms last 14 days**	(6) **male condoms last 14 days**
Treatment	0.169*	-0.135	0.139*	0.005	0.069	-0.099
	(0.088)	(0.132)	(0.079)	(0.124)	(0.066)	(0.119)
Observations	56	56	56	56	56	56
Control mean	0.069	0.552	0.034	0.448	0.034	0.448

Notes: Regressions on the balanced diary sample, N=56. Dependent variables are binary indicators for the use of female condoms (FC) and male condoms (MC). Columns 1 and 2 refer to whether the respondent reports using the method at least once during the full endline period, columns 3 and 4 to whether she reports using it at least once in the last 30 days, and columns 5 and 6 to whether she reports using it at least once in the last 14 days. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Facilitator fixed effects are not included because the sample is small with 57 endline observations, implying that some facilitators perfectly predict the endline outcome variable. Standard errors are bootstrapped with 10,000 replications, clustered at the respondent level. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Impacts on condom use at the sex-act level: We can use the diary data to measure treatment effects at the sex-act level, specifically on the likelihood that a sex act is protected by male or female condoms. We estimate the following linear probability model:

$$\begin{aligned} &Pr\left[Y_{jift}=1|treat_{if},\eta_f,\phi_{if}\right] \\ &=\alpha+\delta\times endline_t+\beta treat_{if}\times endline_t+\eta_f\times endline_t+\phi_{if}, \quad j=1,2,..J \quad t=1,2,..T \end{aligned} \tag{B.28}$$

where Y_{jift} is the outcome variable of interest relating to sex act j, reported by individual i, who was assigned to facilitator f.[32] *endline* is an indicator equal to one if the week t in which sex act j is reported falls in the endline period, i.e. one week or more after programme sessions led by i's facilitator have begun for the treatment group. ϕ_{if} is an individual fixed effect. Standard errors are again clustered at the individual level, since this was the unit of randomisation (Abadie et al., 2017).

Table B.31 shows the results. There is a significant increase in the likelihood that a sex act is protected by a female condom in the full endline period, the last 30 days, and the last 14 days (5.1, 8.5 and 6.7 percentage points respectively, all significant at the 5% level). Consistent with the respondent-level results from the survey data, there is no significant effect on the probability that a sex act is protected by male condoms, and the point estimates are small.

[32] For each sex act and each contraceptive method, we set "sex act full endline period" equal to one if that method was used and the sex act took place at any point during the endline period. Similarly, we set "sex act use last 30 days" equal to one if that method was used and the sex act occurred in the last four weeks prior to a respondent's last diary observation. Finally, we set "sex act last 14 days" equal to one if the method was used and the sex act occurred 14 days prior to a respondent's last diary observation.

Table B.31: Impacts on probability that a sex act by a respondent is protected by male or female condoms – diary subsample

	(1) **female condom full endline**	(2) **male condom full endline**	(3) **female condom last 30 days**	(4) **male condom last 30 days**	(5) **female condom last 14 days**	(6) **male condom last 14 days**
Treat×endline	0.051**	-0.005	0.085**	0.015	0.067**	-0.004
	(0.023)	(0.087)	(0.034)	(0.119)	(0.031)	(0.135)
Facilitator× endline f.e.'s	✓	✓	✓	✓	✓	✓
Observations	349	349	204	204	143	143
Control mean	0.010	0.330	0.009	0.374	0.013	0.387

Notes: Regressions on the balanced diary sample, N=56. Dependent variables are binary indicators for the use of female condoms (FC) and male condoms (MC) by a respondent where the sex act is the unit of observation. Column 1 and 2 report the results for the full endline period, Column 3 and 4 for the last 30 days, and Column 5 nad 6 the last 14 days. All regressions are linear probability individual fixed effects models with the respondent-sex act as the unit of observation (N=349 for the full endline period, N=204 for the last 30 days, and N=143 for the last 14 days). "Treat × endline" is an indicator for observations in the treatment group (i.e. to the first round of the family planning training sessions) during the relevant endline period ("full endline", "last 30 days", or "last 14 days") as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions include facilitator × endline fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$

Table B.32: Likelihood of sex acts: Interaction of treatment and baseline male condom use – diary subsample

	(1) **Sex act per week full endline period**	(2) **Sex act per week last 30 days**	(3) **Sex act per week last 14 days**
Treat × endline	0.139***	0.200***	0.197**
	(0.051)	(0.050)	(0.098)
Treat × endline × male condom at baseline	-0.105	-0.174**	-0.078
	(0.081)	(0.071)	(0.102)
Facilitator × endline f.e.'s	✓	✓	✓
Observations	863	536	367
Control mean	0.469	0.471	0.491

Notes: Regressions on the balanced diary sample, N=56. Dependent variables are binary indicators for whether a respondent had at least one sex act in a particular week. "male condom at baseline" is a binary indicator for whether a respondent reported at baseline that she was currently using male condoms. Column 1 refers to whether the respondent had at least one sex act per week in the full endline period, Column 2 whether she had at least one sex act in the last 30 days, and Column 3 whether she had at least one sex act in the last 14 days. All regressions are linear probability individual fixed effects models with the respondent-week as the unit of observation (N=863 for the full endline period, N=536 for the last 30 days, and N=367 for the last 14 days). "Treat × endline" is an indicator for observations in the treatment group (i.e. to the first round of the family planning training sessions) during the relevant endline period ("full endline", "last 30 days", or "last 14 days") as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions include facilitator × endline fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$

Table B.33: Impacts on proportion of sex acts in a week where the respondent and her partner had discussions about protection – diary subsample

	(1) **Discussion full endline**	(2) **Female-initiated full endline**	(3) **Discussion last 30 days**	(4) **Female-initiated last 30 days**	(5) **Discussion last 14 days**	(6) **Female-initiated last 14 days**
Treat×endline	-0.031	-0.078	-0.126	-0.144*	-0.282**	-0.219***
	(0.111)	(0.078)	(0.103)	(0.075)	(0.110)	(0.064)
Facilitator×endline f.e.'s	✓	✓	✓	✓	✓	✓
Observations	398	398	259	259	179	179
Control mean	0.227	0.192	0.275	0.228	0.311	0.265

Notes: Regressions on the balanced diary sample, N=56. Dependent variables are the proportion of sex acts of a respondent in a particular week where the respondent and her partner had discussion about protection. Column 1 and 2 report the results for the full endline period, Column 3 and 4 for the last 30 days, and Column 5 and 6 the last 14 days. Columns 1, 3, and 5 report the results for any discussion while Columns 2, 4, and 6 report results only for female-initiated discussions. All regressions are linear probability individual fixed effects models with the respondent-week as the unit of observation, and weeks with zero sex acts are counted as missing (N=398 for the full endline period, N=259 for the last 30 days, and N=179 for the last 14 days). "Treat × endline" is an indicator for observations in the treatment group (i.e. to the first round of the family planning training sessions) during the relevant endline period ("full endline", "last 30 days", or "last 14 days") as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions include facilitator × endline fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$

B.5 Cost-Effectiveness and Cost-Benefit Analysis

We estimate the effects on the entire population of Southern Mozambique of scaling up the intervention to cover all women in the age-group typically considered as most sexually active (15-49 years) for the years 2015-30, excluding high-risk groups.[33] We take the current HIV/AIDS national strategic program in Mozambique as given, assuming that commitments including the provision of anti-retroviral therapies (ART) would not change if female condoms were also offered. We first simulate a control projection, where estimates from 2015-16 are taken and projections for 2017-30 are made based on the status quo, with none of the epidemiological and behavioural parameters changed. We then simulate two female condom intervention scenarios, based on the impacts of the intervention estimated from our experiment. In the first scenario, we focus purely on the increase in condom coverage and marginal decrease in average condom effectiveness when individuals adopt female condoms as a result of the intervention. In the second scenario, we also take into account the behavioural response via the estimated increase in the number of sex acts. This second scenario is our preferred estimate, but comparison with the first scenario allows us to quantify the importance of the behavioural response and its negative spillovers.

To model the health impacts of our intervention, we use the we use the AIM module of the SPECTRUM suite of epidemiological models (as used by UNAIDS) to estimate the number of HIV infections and disability-adjusted life years (DALYs) that the scale-up scenarios would help to avert in comparison to the control scenario. Figure B.2 shows the simulated number of new HIV infections per year in the control as well as the two intervention scenarios. Table B.34 summarizes the total number of HIV infections and DALYs that would be averted by 2030.[34]

[33] In the epidemiological model that we use, adults above the median age of first sex are allocated into one of five risk categories, identified for males and females separately. These are: stable couples (men and women reporting a single partner in the last year); multiple partners (men and women with more than one partner in the last year); female sex workers and clients; men who have sex with men; and injecting drug users. Our intervention targets women in the first two categories, whose partners are estimated by the epidemiological model to be primarily in the second category. It does not target individuals in the last three, high-risk categories.

[34] The SPECTRUM suite is developed by Avenir Health, see http://www.avenirhealth.

To estimate the implied financial benefits to the healthcare system, we focus on the reduction in the number of adults and children that require ART, cotrimoxazo (an antibiotic used both to treat and prevent pneumocystis pneumonia and toxoplasmosis in people with HIV/AIDS) and the number of mothers requiring Prevention of Mother-To-Child Transmission for the period from 2015-2030 (see Tables B.35 and B.36 for scenario 1 and 2, respectively).[35] To estimate the cost-savings of our intervention in terms of productivity gains, we estimate the reduction in productivity losses as a result of continued workforce participation of adults who did not get infected with HIV as a result of our intervention.

We next calculate an upper and a lower bound of the intervention costs per participant. For the upper bound, we use the full costs of our intervention as implemented, plus the full cost of acquiring and distributing the subsequent increase in the number of female condoms used between 2015 and 2030, assuming full subsidisation of female condom provision by the government (Columns 3 and 5 in Tables B.39 and B.40 for scenarios 1 and 2, respectively). For the lower bound, we assume that the provision of information about female condoms is included into existing sex education programmes in schools and at health centres. This is a realistic add-on to such programmes, given that they already provide information about and practical demonstrations of male condoms, as well as information about HIV/AIDS and other STIs. The lower bound cost estimates therefore comprise just the costs of acquiring and distributing the additional number of female condoms when adoption subsequently increases, assuming that the government fully subsidises free provision of female condoms (Column 5 in Tables B.39 and B.40 for scenarios 1 and 2, respectively).

Comparing the programme costs in Tables B.39 and B.40 to the DALYs averted in Table B.34allows us to calculate the incremental cost-effectiveness ratio (ICER). This measure is often used to compare the cost-effectiveness of policies across the public health spectrum, in terms of cost per DALY averted (see e.g. (Creese et al., 2002;

org/software-spectrum.php.

[35]Tables B.37and B.38 show the unit costs for counseling, drugs and treatment that are used in the calculations.

Oster, 2005)). Comparing the programme costs to the cost savings in Tables B.35 and B.36 allows us to calculate the internal rate of return (IRR). This is an indicator of cost-benefit, which can be used to evaluate the policy as a financial investment.

In scenario 1 the ICER for the full intervention is -50 USD, i.e. a saving of 50 USD per DALY averted, meaning that scaling up the full intervention is therefore *very cost-effective*.[36] It also offers a positive financial return, with an IRR of 1.02. Meanwhile, the ICER for the lower-cost, add-on intervention is -1,574 USD, i.e. a saving of 1,574 USD. This means that adding female condom provision to existing sex education programs is also *very cost-effective*, and in fact represents a substantial saving per DALY averted compared to the existing set of treatments. It also offers a highly favourable return on investment of 1.82.

In contrast, in scenario 2 the ICER for the full intervention is 7,413 USD, meaning that a full scale-up of the intervention is *not cost-effective*. Nonetheless, the ICER for the lower bound is 3,497 USD, implying that adding female condom provision to existing sex education programs is *cost-effective*. Yet despite being cost-effective in the lower bound scenario, the intervention does not offer a positive financial return on investment: the IRR for the upper-bound cost is 0.21 and for the lower-bound cost is 0.36.

In summary, in scenario two when taking account the observed increase in risky sex acts, only adding female condom provision to existing sex education programmes is *cost-effective*. However, there are still several reasons to believe that our estimates of the IRR and ICER are conservative, and thus that scale-up of both the full programme and adding female condoms to existing initiatives could be substantially more cost-effective than we estimate. First, we use an upper bound for the estimated costs of condoms, which is likely to be highly conservative given that the scale-up of the intervention to the entire female population of South Mozambique would lead to economies of scale in production and procurement. Second, as mentioned above, potentially sizeable benefits

[36]Following the recommendations of the Commission on Macroeconomics and Health, WHO-CHOICE deems interventions *highly cost-effective* if the ICER is less than GDP per capita, cost-effective if the ICER is between one and three times GDP per capita, or *not cost-effective* if the ICER is higher than three times GDP per capita (Walensky et al., 2013). The GDP per capita of Mozambique was 511 USD in 2014.

such as reduction in unwanted pregnancies and other STIs, indirect costs to the health system, and costs for orphan care, are not included in our estimates.

Figure B.2: Simulation of annual number of HIV infections

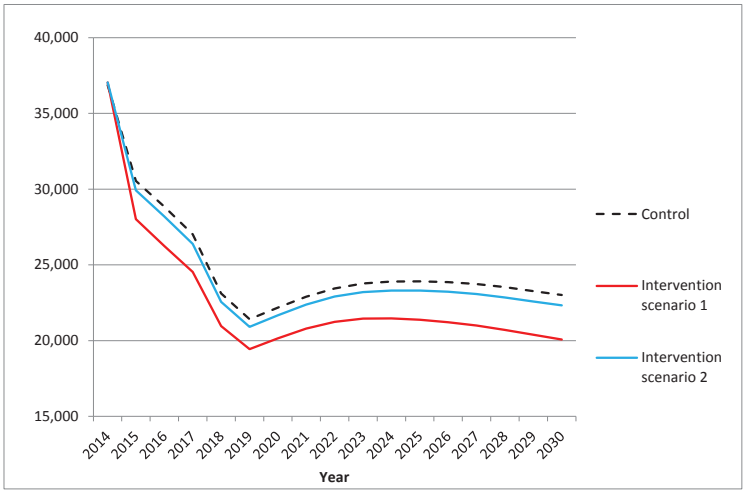


Table B.34: Simulation of impact on HIV infections and DALYs averted by 2030

	# HIV infections averted	**# DALYs averted**
Scenario 1: condom use response only	39,425	72,628
Scenario 2: condom use response & sex act response	9,647	3,607

Notes: Results from simulations based on 2017 UNAIDS data of South Mozambique using the DemProj, AIM, and GOALS module of Avenir Health's SPECTRUM software. Total population (15-49 years) in 2014 was 3,048,905. Columns 1 and 2 present the number of HIV infections and the number of Disability-Adjusted Life Years (DALYs) averted in each scenario, respectively. The statistics are calculated by comparing control projections up to 2030 without any changes to the demographic and behavioural data (control) with intervention projections where behavioural data (condom use) and epidemiological data (condom efficacy) are changed from 2015 onward.

Table B.35: Cost savings ART, cotrimoxazol, and PMTCT – Scenario 1

	Reduction in population on:				Total cost savings (USD)			
Year	Adult Art	Child ART	Cotrimoxazol	PMTCT	Adult ART	Child ART	Cotrimoxazol	PMTCT
2015	-174	96	110	24	-103,091	16,530	3,054	3,271
2016	1,281	231	371	214	780,727	40,948	10,659	30,039
2017	4,035	318	660	397	2,533,600	58,061	19,519	57,398
2018	5,280	406	929	566	3,414,407	76,352	28,318	84,287
2019	6,611	430	1,110	711	4,403,084	83,291	34,845	109,056
2020	8,002	524	1,211	835	5,489,928	104,544	39,147	131,918
2021	9,445	697	1,304	961	6,673,845	143,231	43,408	156,379
2022	10,932	811	1,410	1,082	7,956,973	171,657	48,340	181,351
2023	12,462	931	1,568	1,195	9,342,269	202,968	55,385	206,299
2024	14,022	1,058	1,699	1,303	10,827,613	237,575	61,800	231,692
2025	15,601	1,197	1,824	1,405	12,408,144	276,852	68,332	257,324
2026	17,188	1,342	1,940	1,500	14,080,152	319,700	74,862	282,965
2027	18,781	1,491	2,048	1,590	15,846,798	365,852	81,432	308,941
2028	20,361	1,644	2,148	1,673	17,695,472	415,496	87,940	334,820
2029	21,926	1,754	2,240	1,749	19,627,422	456,595	94,458	360,531
2030	23,524	1,788	2,283	1,819	21,689,329	479,410	99,194	386,210
TOTAL	**189,278**	**14,718**	**22,854**	**17,024**	**152,666,673**	**3,449,062**	**850,692**	**3,122,483**

Notes: UNAIDS data 2017. The reduction in the population on Antiretroviral therapy (ART), Cotrimoxazol (an antibiotic used both to treat and prevent pneumocystis pneumonia and toxoplasmosis in people with HIV/AIDS) and Prevention of mother-to-child transmission (PMTCT) is calculated by making projections in the control and intervention scenario of the number of individuals needing ART, Cotrimoxazol and PMTCT. The number of people receiving these drugs are then calculated by multiplying the number of individuals needing ART, Cotrimoxazol and PMTCT to the coverage levels of the 2015–2019 national strategic HIV/AIDS plan. To calculate the total cost, the difference in the population on ART, Cotrimoxazol and PMTCT is then multiplied by the discounted cost for Adult ART (592 USD in 2014, discounted at 3%), Child ART (172 USD in 2014, discounted at 3%), and PMTCT (136 USD in 2014, discounted at 3%).

Table B.36: Cost savings ART, cotrimoxazol, and PMTCT – Scenario 2

	Reduction in population on:				Total cost savings (USD)			
Year	Adult Art	Child ART	Cotrimoxazol	PMTCT	Adult ART	Child ART	Cotrimoxazol	PMTCT
2015	-425	-18	-51	-46	-251,485	-3,072	-1,416	-6,269
2016	-218	8	23	8	-132,870	1,418	668	1,123
2017	421	30	108	60	264,496	5,477	3,190	8,675
2018	751	53	186	110	485,911	9,967	5,682	16,381
2019	1,099	61	245	153	732,116	11,816	7,685	23,468
2020	1,460	84	280	190	1,001,574	16,759	9,061	30,017
2021	1,834	128	311	227	1,295,727	26,304	10,354	36,939
2022	2,218	158	343	262	1,614,198	33,442	11,766	43,913
2023	2,609	192	388	294	1,955,971	41,858	13,692	50,755
2024	3,007	227	425	324	2,321,816	50,973	15,457	57,612
2025	3,409	268	459	353	2,710,954	61,985	17,206	64,652
2026	3,808	310	491	379	3,119,752	73,850	18,962	71,496
2027	4,207	354	520	403	3,550,091	86,862	20,692	78,304
2028	4,600	402	546	426	3,997,793	101,599	22,359	85,256
2029	4,987	434	570	445	4,464,507	112,977	24,045	91,730
2030	5,380	444	581	463	4,960,420	119,048	25,242	98,304
TOTAL	**39,148**	**3,135**	**5,427**	**4,051**	**32,090,970**	**751,265**	**204,644**	**752,355**

Notes: UNAIDS data 2017. The reduction in the population on Antiretroviral therapy (ART), Cotrimoxazol (an antibiotic used both to treat and prevent pneumocystis pneumonia and toxoplasmosis in people with HIV/AIDS) and Prevention of mother-to-child transmission (PMTCT) is calculated by making projections in the control and intervention scenario of the number of individuals needing ART, Cotrimoxazol and PMTCT. The number of people receiving these drugs are then calculated by multiplying the number of individuals needing ART, Cotrimoxazol and PMTCT to the coverage levels of the 2015–2019 national strategic HIV/AIDS plan. To calculate the total cost, the difference in the population on ART, Cotrimoxazol and PMTCT is then multiplied by the discounted cost for Adult ART (592 USD in 2014, discounted at 3%), Child ART (172 USD in 2014, discounted at 3%), and PMTCT (136 USD in 2014, discounted at 3%).

Table B.37: Unit costs ART for adults and children Mozambique 2015

Adults (costs per patient per year)	
First line ART drugs	124.52
Second line ART drugs	327.17
Laboratory costs for ART treatment	76.31
Cotrimoxazole prophylaxis	27.88
TB prophylaxis	1.26
Nutrition supplements in first six months	17.30
Children (costs per patient per year)	
Children ART drugs	124.48
Children Laboratory costs for ART treatment	29.00
Service delivery costs	
Cost per in-patient day	0.00
Cost per out-patient visit	16.62
Service delivery requirements (per patient per year)	
ART: in-patient days	0.00
ART: out-patient visits	1.00
OI treatment: in-patient days	0.00
OI treatment: out-patient visits	1.00

Notes: Data for South Mozambique 2015. Based on the SPECTRUM Resource Needs Model. ART: Antiretroviral therapy, TB: Tuberculosis, OI: Opportunistic infections.

Table B.38: Unit cost PMTCT for mothers Mozambique 2015

Counseling (per mother)	
Pre-test	3.97
Post-test for HIV+ mothers	47.35
Post-natal for HIV+ mothers	53.1
HIV testing (per test)	
Mother	11
PCR test for infant after birth	5.9
Infant after cessation of breastfeeding	1.3
ARVs (cost per person per day)	
Nevirapine (NVP), 200mg for mother	0.81
Nevirapine (NVP), for infant	0.003
Azidothymidine (AZT)	0.45
Lamivudine (3TC)	0.43
Triple treatment (AZT+3TC+NVP/Efavirenz (EFV))	0.45
Triple prophylaxis	0
Service delivery (per mother)	11.52

Notes: Data for South Mozambique 2015. Based on the SPECTRUM Resource Needs Model. PMTCT: Prevention of mother-to-child transmission, PCR: Polymerase chain reaction test (to diagnose HIV infection in infants, ARVs: Antiretroviral drugs.

Table B.39: Programme and condom unit and distribution costs – Scenario 1

(1) Year	(2) Female population to be treated	(3) Programme costs intervention (USD)	(4) # of additional condoms	(5) Cost additional condoms (USD)
2015	1,653,100	47,774,590	11,515,027	5,181,762
2016	83,605	2,488,670	11,802,213	5,470,326
2017	85,271	2,614,410	12,089,754	5,771,709
2018	86,933	2,745,328	12,382,153	6,088,641
2019	88,931	2,892,677	12,682,188	6,423,261
2020	90,935	3,046,597	12,988,868	6,775,946
2021	92,599	3,195,417	13,302,829	7,147,923
2022	95,110	3,380,528	13,632,923	7,545,049
2023	96,410	3,529,537	13,971,767	7,964,558
2024	98,123	3,700,017	14,321,137	8,408,626
2025	99,651	3,870,363	14,677,351	8,876,310
2026	100,845	4,034,240	15,035,168	9,365,485
2027	102,920	4,240,766	15,398,190	9,879,361
2028	103,969	4,412,509	15,743,360	10,403,845
2029	105,465	4,610,280	16,094,980	10,955,296
2030	106,651	4,801,989	16,450,994	11,533,551
TOTAL	**3,090,518**	**101,337,918**	**222,088,903**	**127,791,649**

Notes: UNAIDS data 2017. Column 1 shows the female population to be treated. In 2015 this is the entire sexually active population (age 15-49 years). From 2016 to 2030, only female 15 year olds and female migrants are treated. The programme costs of the intervention in Column 2 are calculated by multiplying the total discounted cost per person of Pathfinder's programme (28.90 USD in 2015, discounted at 3% per year) with the female population to be treated. The number of additional condoms required in Column 3 are calculated by combining data on the population, # of partners per risk group, # of sex acts, condom wastage, % of condom use in the intervention and control, and the condom efficacy in intervention and control. The cost of condoms in Column 4 is calculated by multiplying the discounted unit and distribution costs of female condoms (0.45 USD in 2015, discounted at 3% per year) with the # of additional condoms required.

Table B.40: Programme and condom unit and distribution costs – Scenario 2

(1) Year	(2) Female population to be treated	(3) Programme costs intervention (USD)	(4) # of additional condoms	(5) Cost additional condoms (USD)
2015	1,653,100	47,774,590.00	12,802,329.56	5,761,048.30
2016	83,605	2,488,670.04	13,121,611.50	6,081,866.93
2017	85,271	2,614,409.71	13,441,288.47	6,416,938.32
2018	86,933	2,745,327.65	13,766,391.66	6,769,308.54
2019	88,931	2,892,676.83	14,100,017.75	7,141,362.39
2020	90,934	3,046,563.69	14,441,046.81	7,533,509.03
2021	92,599	3,195,416.60	14,790,174.37	7,947,108.75
2022	95,109	3,380,492.92	15,157,229.11	8,388,665.98
2023	96,410	3,529,536.87	15,533,996.65	8,855,100.99
2024	98,123	3,700,016.69	15,922,443.58	9,348,829.83
2025	99,651	3,870,363.46	16,318,471.64	9,868,797.60
2026	100,844	4,034,199.57	16,716,247.56	10,412,638.23
2027	102,920	4,240,766.07	17,119,772.64	10,983,916.00
2028	103,969	4,412,509.27	17,503,412.21	11,566,957.92
2029	105,464	4,610,236.74	17,894,181.00	12,179,949.44
2030	106,645	4,801,718.68	18,289,792.97	12,822,705.68
TOTAL	**3,090,508**	**101,337,495**	**246,918,407**	**142,078,704**

Notes: UNAIDS data 2017. Column 1 shows the female population to be treated. In 2015 this is the entire sexually active population (age 15-49 years). From 2016 to 2030, only female 15 year olds and female migrants are treated. The programme costs of the intervention in Column 2 are calculated by multiplying the total discounted cost per person of Pathfinder's programme (28.90 USD in 2015, discounted at 3% per year) with the female population to be treated. The number of additional condoms required in Column 3 are calculated by combining data on the population, # of partners per risk group, # of sex acts, condom wastage, % of condom use in the intervention and control, and the condom efficacy in intervention and control. The cost of condoms in Column 4 is calculated by multiplying the discounted unit and distribution costs of female condoms (0.45 USD in 2015, discounted at 3% per year) with the # of additional condoms required.

B.6 Linear versus nonlinear regression specification

The regression equations in the main body of the paper are ANCOVA (Analysis of Covariance) linear probability models (LPM) of the following form:

$$Pr\left[Y_{if1} = 1 | Y_{if0}, treat_{if}, \eta_f\right] = \alpha + \delta Y_{if0} + \beta treat_{if} + \eta_f, \tag{B.29}$$

where Y_{if1} is the outcome variable of interest at endline, and Y_{if0} is its value at baseline. The variable $treat_{if}$ is a dummy for being assigned to the treatment group, i.e. to receiving the programme in the first rather than the second phase. The parameter β represents the intent-to-treat effect. The parameter η_f is a facilitator fixed effect, which is included for inference since randomisation was blocked on the seventeen facilitators (Bruhn and McKenzie, 2009). Standard errors are robust to individual-level heteroskedasticity, as this was the level of randomisation (Abadie et al., 2017).

Table A.1 in the main paper displays the main treatment impacts of the intervention. Since the majority of outcome variables are binary indicators for whether a respondent has used a certain STI protection technology or had a sex act in a particular week, the preferred specification could have been a logit. Two reasons in favour of logit are the fact that LPM tends to produce consistently biased and inconsistent estimates of structural parameters and marginal effects.[37] For maximum likelihood estimation of the logit model both the marginal effect and the standard deviation decrease as sample sizes increase (Amemiya, 1977; Horrace and Oaxaca, 2006).[38] However, estimating marginal effects with the wrong nonlinear model can also produce inconsistent estimates. This is especially true in our case because maximum likelihood estimators (MLEs) such as logit are well-known to suffer from small-sample bias and underestimation of the true probability of rare events (Leitgöb). The degree of bias is strongly dependent on the number of cases in the less frequent of the two categories in the dependent variable. Even with a large sample, if there are a few cells with very few observations this is

[37] The marginal effect implied by LPM remains consistently biased even if sample sizes increase, while the inconsistency in LPM is driven by asymptotic bias

[38] http://davegiles.blogspot.co.uk/2012/06/another-gripe-about-linear-probability.html

problematic (King and Zeng, 2001a,b, 2002). Only 2% of women in our sample report they are currently using female condoms at baseline. With intent-to-treat effects of 5-20 percentage points female condom use remains a relatively rare event. This problem is exacerbated by the fact that we require fixed effects for facilitators in the estimating equation, as these were used for stratification (Bruhn and McKenzie, 2009). If the stratification dummy is used, the number of observations per cell is substantially reduced, and we run into the problem that facilitators perfectly predict outcome variables, and estimating a treatment effect becomes impossible because there is no variation between base and endline and treatment and control. This is true regardless of the specification (LPM, logit, or probit) but LPM doesn't drop these observations, while logit does and correctly so. The resulting variation in sample sizes for different dependent variables, while using the same regression specification, creates uncertainty about the comparison of effects.

One way to correct for this bias is to use the Bias Correction Method proposed by King and Zeng (2001a) through re-estimating our results with their 'rare event logit' (relogit) estimator. Alternatively, penalized maximum likelihood estimators can be used, such as firthlogit or exact logistic regression because they deal specifically with concerns about bias due to small samples and 'separation' in logistic regression (Firth, 1993; Heinze and Schemper, 2002).[39] These estimators have the attraction of producing finite, consistent estimates of regression parameters when the maximum likelihood estimates do not even exist because of complete or quasi-complete separation.

There is a trade-off between using the correction method and the penalized estimators. The correction method by (King and Zeng, 2001a) overcorrects bias in MLEs as the sample size gets small but it does allow for clustering of standard errors (Leitgöb). Firthlogit and exact logistic regression seem unbiased and converge, even in the case of low numbers of observations when stratification dummies are used (Heinze and Schemper, 2002). A complication, however, is that it does not allow clustering of standard errors that are robust to individual level heteroskedasticity. This is required

[39] A condition in the data in which maximum likelihood estimates become inestimable because they tend to infinity

because this was the level of randomization (Abadie et al., 2017).

To demonstrate robustness of our results to the use of nonlinear specifications, the results are re-estimated with nonlinear models. Tables A.1, B.41, B.42, and B.43 show the estimates from the LPM, logit, Firthlogit, and relogit estimators for the main treatment effects in the survey, respectively. The sample sizes in the logit specification vary across specifications, but are consistent for the LPM, Firthlogit, and relogit. Qualitatively the estimates for the LPM specification, and the estimated marginal effects for the logit and Firthlogit are the same, although the LPM and Firthlogit estimates appear more consistent with each other. The predicted marginal effects for the treatment effect on ever use, last 30 day use, and current use are 26.3, 11.7, and 6.7 percentage points respectively, and are significant at the one percent level. Tables A.3 and B.44 show the panel fixed effects specifications for the LPM and logit, respectively. These estimates are also consistent. The weekly nature of the diary data implies that we have multiple observations per respondent, and don't face the perfect predictor problem when we re-estimate the panel fixed effects LPM model with a logit model. Nonlinear panel models do suffer from the incidental parameter problem which causes parameters to not be point-identified (Heckman, 1981; Lancaster, 2002; Chernozhukov et al., 2013).

Small sample sizes do become a problem when we estimate treatment impacts conditional on the level of baseline bargaining power. Due to the perfect predictor problem MLE logit estimations don't converge. Firthlogit and relogit overcome these problems. Relogit does so by estimating the same model as standard logistic regression, but correcting the estimates for the fact that in rare events data values of one are more statistically informative than values of zero. A correction of the weighting of these values can be used to reduce the variance and make the parameter estimates more informative (King and Zeng, 2001a). Firthlogit, rather then correcting the estimate ex-post, makes a systematic correction in the mechanism that produces the maximum likelihood estimate, namely the score equation, thereby not necessarily requiring the existence of a finite estimate, making the approach especially suitable for small samples (Firth, 1993). Firthlogit, however, doesn't allow for clustering of standard errors at the level of ran-

domisation. The lack of clustering produces insignificant estimates, although still of the same sign as the LPM. If we re-estimate the LPM without clustering of standard errors we produce qualitatively similar results as with the Firthlogit. Fortunately relogit does allow for clustering and the results are qualitatively similar to the estimates from the LPM in terms of direction and significance. For ease of interpretation the predicted marginal heterogenous treatment effect on the current use of female condoms by the five bargaining power indices is presented in Figure 3. These predicted marginal effects are based on the estimates from the rare events logit (see Table B.45). The dashed line is the average "intent-to-treat" effect. The solid lines represents the predicted margin of the treatment effect, estimated from the interaction of treatment with the bargaining power indices, with 95% confidence intervals. For all bargaining power principal components the treatment is significantly higher for those respondents with low female bargaining power, as predicted by our model.

Table B.41: Main treatment outcomes – Logit specification

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	0.275***	0.003	0.011	0.129	-0.050	0.165**	0.064	0.030
	(0.068)	(0.045)	(0.045)	(0.081)	(0.057)	(0.077)	(0.059)	(0.054)
Facilitator f.e.'s	✓	✓	✓	✓	✓	✓	✓	✓
Observations	172	193	218	112	227	141	227	227
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Maximum likelihood estimation of the logit model on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill, injectables or IUD. The reduced number of observations are the result of the fact that facilitators perfectly predict outcome variables, and estimating a treatment effect becomes impossible because there is no variation between base and endline and treatment and control. Logit correctly drops such observations. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days (this was only asked for condoms, not for other contraceptive methods), and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are logit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.42: Main treatment outcomes – Firthlogit specification

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	0.198***	-0.001	0.012	0.074*	-0.049	0.105***	0.060	0.029
	(0.043)	(0.047)	(0.048)	(0.040)	(0.058)	(0.040)	(0.058)	(0.054)
Facilitator f.e.'s	✓	✓	✓	✓	✓	✓	✓	✓
Observations	227	227	227	227	227	227	227	227
Control mean endline	0.088	0.824	0.735	0.010	0.363	0.020	0.353	0.412

Notes: Penalized maximum likelihood estimation of the Firthlogit model on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill, injectables or IUD. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days (this was only asked for condoms, not for other contraceptive methods), and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are Firthlogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.43: Main treatment outcomes – relogit specification

	(1) **Ever use female condoms**	(2) **Ever use male condoms**	(3) **Ever use other**	(4) **Last 30 days use female condoms**	(5) **Last 30 days use male condoms**	(6) **Current use female condoms**	(7) **Current use male condoms**	(8) **Current use other**
Treatment	1.565***	-0.182	0.128	1.496	-0.355	1.616**	0.250	0.126
	(0.465)	(0.400)	(0.408)	(1.062)	(0.297)	(0.762)	(0.289)	(0.340)
Facilitator f.e.'s	✓	✓	✓	✓	✓	✓	✓	✓
Observations	227	227	227	227	227	227	227	227
Control mean endline	0.088	0.824	0.735	0.412	0.363	0.020	0.353	0.412

Notes: Maximum likelihood estimation of the "rare events logit (relogit)" model on the balanced sample, N=227. Dependent variables are binary indicators for the use of female condoms (FC), male condoms (MC) and other modern contraceptive methods (other), such as the pill, injectables or IUD. Columns 1-3 refer to whether the respondent has ever used the method, columns 4 and 5 to whether she has used it in the last 30 days (this was only asked for condoms, not for other contraceptive methods), and columns 6-8 whether she is currently using it. "Treatment" is an indicator for being assigned to the treatment group (i.e. to the first round of the family planning training sessions) as opposed to the control group (i.e. the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. All regressions are relogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator indicators (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.44: Impacts on likelihood of sex acts per respondent week – Logit

	(1) **# Sex acts per week full endline period**	(2) **# Sex acts per week last 30 days**	(3) **# Sex acts per week last 14 days**
Treat×endline	0.131**	0.168**	0.142*
	(0.054)	(0.068)	(0.084)
Observations	863	536	367
Control mean	0.469	0.471	0.491

Notes: Maximum likelihood estimation of the logit model on the balanced diary sample, N=56. Dependent variables are binary indicators for whether a respondent had at least one sex act in a particular week. Column 1 refers to whether the respondent had at least one sex act per week in the full endline period, Column 2 whether she had at least one sex act in the last 30 days, and Column 3 whether she had at least one sex act in the last 14 days. All regressions are logit fixed effects models with the respondent-week as the unit of observation (N=863 for the full endline period, N=536 for the last 30 days, and N=367 for the last 14 days). "Treat × endline" is an indicator for observations in the treatment group (i.e. respondents assigned to the first round of the family planning training sessions) during the relevant endline period ("full endline", "last 30 days", or "last 14 days") as opposed to the control group (i.e. respondents assigned to the second round of training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions include facilitator × endline fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors (in parentheses) are robust to individual-level heteroskedasticity, since this was the level of randomisation. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.

Table B.45: Impacts on current use of female condoms – heterogeneity by bargaining power – Relogit

	(1) Current use female condoms	(2) Current use female condoms	(3) Current use female condoms	(4) Current use female condoms	(5) Current use female condoms
Treatment	28471757.599***	5.063***	21.857***	1.540	6.300**
	(0.763)	(0.850)	(0.983)	(1.218)	(2.859)
Assets score 1	70020910.740***				
	(0.791)				
Treatment*Assets score 1	-7.002e+07				
	(0.000)				
Assets score 2		23.603***			
		(1.523)			
Treatment*Assets score 2		-20.773***			
		(1.753)			
Assets score 3			21.840***		
			(0.404)		
Treatment*Assets score 3			-22.168***		
			(0.593)		
Decision-making				0.452	
				(0.519)	
Treatment*Decision-making				-1.500**	
				(0.669)	
Power dynamics					1.308*
					(0.675)
Treatment*Power dynamics					-1.590**
					(0.765)
Controls	✓	✓	✓	✓	✓
Observations	201	201	201	182	182
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Maximum likelihood estimation of the "rare events logit (relogit)" model on the balanced sample of respondents (N=227) with a stable relationship at baseline (N=206); N=201 of them have non-missing values on all assets and control variables; N=182 of them have non-missing values on all decision-making, power dynamics, and control variables. Dependent variables are binary indicators for current use of female condoms (FC). "Treatment" is a dummy for being assigned to the treatment group (i.e. to the first round of the family planning training sessions). Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as described in Table 3 and identified in Table B.8. "Decision-making" and "Power dynamics" are the first two principal components from all the survey questions referring to these two modules, as identified in Table B.9. For comparability the components are scaled so that the woman with least bargaining power on that measure has a score of zero. The components are normalized such that a one point increase in each measure represents an increase of one standard deviation. All regressions are relogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator indicators due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation. Controls are: respondent's age, education, and income in the last 30 days; whether the respondent has a job, is married or in a stable relationship, and whether the respondent is the household head. Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$.