The Power to Protect: Household Bargaining and Female Condom Use*

Rachel Cassidy[†], Marije Groot Bruinderink[‡], Wendy Janssens[§] & Karlijn Morsink[¶]

January 25, 2019

Abstract: Women may face systematically greater benefits than men from adopting certain technologies. Yet women commonly hold lower bargaining power, which may constrain adoption when decisions are joint. Introducing a version of the technology that is second-best in terms of cost or effectiveness, but more acceptable to men, may increase adoption and welfare. We conduct a field experiment introducing female condoms – which are less effective and more expensive than male condoms, but often preferred by men – in a setting with high HIV prevalence. We find strongest adoption among women with low bargaining power, who were previously having unprotected sex.

JEL classification: C78, O33, C93, J16, I12

*We thank Dan Anderberg, Nava Ashraf, Martin Browning, James Fenske, Glenn Harrison, Eline Korenromp, Emily Oster, Simon Quinn and Chris Woodruff for very helpful comments. We also thank audiences at seminars and conferences including NEUDC, the Royal Economic Society, the Royal Econometric Society, ECBE, and EUDN. We are grateful to Lene Boehnke and Emilie Berkhout for excellent research assistance, and to Balthazar Chilundo from Eduardo Modlane University, Pathfinder, Oxfam Novib, and WeConsult for research support. This research project was funded by the Netherlands Organisation for Scientific Research (NWO) WOTRO Science for Global Development (grant no. W 07.40.203) and the Universal Access to Female Condoms (UAFC) Joint Programme (grant no. A-02974-02-01/506671). IRB approval was granted by the National Research Ethics Committee of Mozambique, the Comité Institucional de Bioética em Saúde da Faculdade de Medicina/Hospital Central de Maputo, Mozambique.

[†]Institute for Fiscal Studies & University of Oxford

[‡]AIGHD

[§]AIGHD, Tinbergen Institute & VU Amsterdam

[¶]Utrecht University & University of Oxford. 120 Walton St, Oxford OX2 6GG, UK, Corresponding author. Email: karlijn.morsink@economics.ox.ac.uk

1 Introduction

The costs and benefits of adopting household technologies may differ systematically across genders. There is evidence that women have a stronger preference for risk reduction (Agnew et al., 2008), investment in children's education (Duflo, 2003), and investment in health via nutritious food (Duflo and Udry, 2004; Attanasio and Lechene, 2014). Women may also bear more of the costs of technology non-adoption, through responsibility for domestic chores, caring duties, and greater exposure to certain health and safety risks. If a technology can be adopted unilaterally then even women with low bargaining power may be able to adopt, for example the pill or concealable contraceptives (Goldin and Katz, 2002; Ashraf et al., 2014b). When adoption of a technology requires agreement between partners, intra-household bargaining matters, and men's preferences may constrain household adoption. Examples include improved cookstoves (Miller and Mobarak, 2013), private latrines, anti-malarial bednets, and condoms. One way to increase adoption may be to target men's preferences directly (Stopnitzky, 2017), although this can prove difficult (Creese et al., 2002). Another option is to increase women's bargaining power (Bandiera et al., 2015; Ashraf et al., 2017), although this may be hard to achieve at scale without broader changes in labour and marriage markets. In contexts where low female bargaining power and male preferences continue to constrain adoption, a second-best solution may be to introduce a variant of the technology that is more acceptable to men, even if less effective or more costly.

This paper puts forward and tests this idea, using a field experiment in Maputo Province, Mozambique. We study adoption of condoms: a technology which is observable to both parties and hence requires joint adoption;[1] but where women face higher costs from non-adoption, via higher risk of contracting HIV (in this context) and unwanted pregnancy.[2] We examine how intra-household bargaining affects adoption of female condoms when they are introduced, in a setting where only male condoms are widely available. Female condoms are second-best insofar as they have lower effective-

[1]Female condoms can be inserted by women prior to intercourse, but remain visible.

[2]In 2015, women accounted for 59% of all individuals aged 15 and over living with HIV in Sub-Saharan Africa, and the rate of new infections among young women aged 15-24 was double that among young men (UNAIDS, 2016a). Reasons for this gender disparity include that women tend to have older partners, lower access to sexual and reproductive health services, and a higher biological risk than men of becoming infected from heterosexual intercourse (UNAIDS, 2016b).

ness than male condoms and higher unit cost.[3] However, female condoms are viewed by men as less uncomfortable or stigmatising than male condoms (Philpott et al., 2006; Wanyenze et al., 2011; Koster et al., 2015). We show that women with lower bargaining power — many of whom are unable to convince their partners to use male condoms at baseline — convince their partners to adopt female condoms when they are made freely available. An illustrative cost-benefit analysis shows that this could lead free provision of female condoms to be cost-effective, but that this result is sensitive to the behavioural response that we also observe: namely, an increase in the number of sex acts.

Condoms are an important technology from a public health perspective, as they are the only well-established protection against HIV/AIDS and other sexually transmitted infections (STIs). They exemplify technologies where adoption is partially or fully observable within the household, agreement of both partners is needed to ensure sustained and proper use, and hence bargaining may constrain adoption. Indeed there is evidence that women may struggle to convince their partners to use male condoms (Anderson, 2018), helping to explain their persistent under-adoption.[4] Condoms are also a particularly good technology for studying the implications of introducing a second-best version. This is because existing epidemiological models of HIV transmission allow us to quantify the potential trade-offs between improving condom coverage and decreasing average effectiveness — as well as behavioural responses such as increases in the frequency of risky sex — while taking into account the negative externalities from HIV transmission.

We evaluate a condom programme in the slums of Maputo, Mozambique. The programme seeks to increase condom use by offering female condoms alongside male condoms. Women attend a series of group sessions that provide information about contraceptives including female condoms. Female condoms are also added to the set

[3]Female condoms have 95% efficacy at preventing pregnancy in the first year of use if perfectly used, or 79% effectiveness in ordinary use (Farr et al., 1994; Trussell, 2011; Beksinska et al., 2012). Male condoms are 98% effective if perfectly used, or 85% effective in ordinary use (`https://www.who.int/news-room/fact-sheets/detail/family-planning-contraception`). The unit production cost for female condoms at current volumes is $0.57 compared to $0.03 for male condoms (Mantell et al., 2015). There is currently a monopoly on the production of WHO-approved female condoms, and consequent low production volumes (Peters et al., 2010). Lower-cost female condoms have been developed in India and approved by the EU, but are still awaiting WHO approval (*ibid.*). Costs would likely substantially decrease at a larger scale of production (Dowdy et al., 2006).

[4]An estimated 3.3 billion risky sex acts took place without condoms in Sub-Saharan Africa in 2015, leading to 910,000 new HIV infections (UNAIDS, 2016a).

of products carried by local health workers — which already includes male condoms — that participants can access freely and discreetly at the end of each session. The intervention thus allows us to study which women, if any, adopt female condoms when informational, access, and price constraints are alleviated. Importantly, free provision allows us to study couples' willingness to adopt unconfounded by their ability to pay, which may be correlated with female bargaining power. Free provision is also arguably the most relevant policy option in countries with high HIV/AIDS prevalence, where male condoms are typically already provided for free by the government.

We conduct a phased-in randomised control trial to assess the short-run impacts of the programme on women who were assigned to receive it at the end of 2014, compared to those who were assigned to receive it six months later. In addition to baseline and endline data, we collect weekly sexual diary data for a subsample of the women. This allows us to investigate impacts at the sex-act level, including effects on the frequency of sex acts. It also reduces concerns about recall bias and misreporting, which we further mitigate by recording the number of condoms that all participants take with them after each session. To measure bargaining power, we collect information about assets brought by the woman to the relationship, and also use two different survey modules covering decision-making and power dynamics in the relationship (Donald et al., 2017).

To formalise our predictions, we introduce a collective model of the household where partners jointly choose STI protection methods. Both men and women value the levels of pleasure and of health protection associated with different contraceptive technologies. However, for the reasons outlined above, we argue that the marginal rate of substitution between health and pleasure is greater for women than for men. When the only STI protection options available are male condoms or unprotected sex,[5] the model predicts that women on average prefer to use male condoms, but that those with low bargaining power are unable to convince their partners to do so. When female condoms are introduced, the model predicts two effects. First, on the intensive margin, women with intermediate bargaining power adopt female condoms. This is because some women with low bargaining power who were previously having unprotected sex are now able to convince their partners to adopt female condoms (but not male condoms), increasing condom coverage. At the same time, some women with intermediate bargaining power

[5] This includes sex protected by pure contraceptives such as the pill, but not by an STI protection method; see Section 3 for details.

who were previously using male condoms also substitute into using female condoms, decreasing average condom effectiveness. The relative magnitudes of each of these margins of response are important to determine empirically, in order to establish total effects on transmission of HIV/AIDS and other STIs. Second, on the extensive margin, some couples who were not previously having sex now have sex with female condoms.

The results show a large impact of treatment on female condom use: an increase of 18.4 percentage points in the proportion of women who have ever used female condoms (equivalent to 209% of the endline mean in the control group) and an increase of 7.7 percentage points (385%) in the proportion of those currently using female condoms. We see no significant evidence of substitution away from male condoms, which is in line with our model's predictions for a sample where few women have high bargaining power. As predicted by the model, adoption of female condoms is driven by women with lower baseline bargaining power, and those who are having unprotected sex at baseline. These results are robust across our measures of bargaining power. On the extensive margin, the diary data show that treatment leads to an increase of 9.1 percentage points (19%) in the probability that an individual has sex each week . We rule out various alternative explanations for the heterogeneous treatment effect by bargaining power: this result is robust to controlling for distance to health centres (proxying access to condoms), use of other contraceptives, HIV status, and beliefs about HIV risk. Nor do we see impacts of our treatment on HIV knowledge, or on our measures of bargaining power themselves.

Given that this is a second-best technology, a social planner should weigh the observed increase in condom coverage against the reduction in average condom effectiveness and increase in production and distribution costs.[6] To demonstrate the potential magnitude of this trade-off, we conduct an illustrative exercise in which we estimate the costs and benefits of scaling up access to female condoms to all of Southern Mozambique, focusing solely on the benefits in terms of reduced HIV transmission. In our naïve scenario before accounting for the behavioural response (i.e. the observed increase in sex acts), both our full programme and adding female condoms to existing sex education programmes actually imply a cost saving. Intuitively, this is because low female bar-

[6] Given that the negative health effects and externalities of unprotected sex are large in the context of our study, it is reasonable to assume that these are the social planner's first-order concern, and to abstract from quantifying individuals' pleasure from using different types of condoms.

gaining power implies that the main margin of female condom adoption is from women previously having unprotected sex, rather than substitution away from male condoms. However, once we account for the increase in sex acts, only adding female condoms to existing sex education programmes has the potential to be cost-effective. These illustrative simulations thus show how accounting for behavioural responses that may offset direct impacts of a programme is crucial (Greenwood et al., 2017).

In terms of the literature on contraceptive technologies, to our knowledge this is the first experimental study explicitly to consider how intra-household bargaining may constrain adoption of male condoms.The existing literature on bargaining within couples focuses on fertility (Eswaran, 2002), and emphasises limited commitment or imperfect information (Rasul, 2008; Ashraf et al., 2014b). In contrast, we emphasise bargaining over STI protection, where use of the technology is fully observable and potentially negotiated each time. Gertler et al. (2005) model bargaining over male condom use between female sex workers and male clients in Mexico, as a finite-horizon, non-cooperative interaction mediated by price. Our contribution is to model bargaining over condoms within the collective model, capturing the efficiency arising from the repeated household bargaining process that takes place within couples.

Our study also highlights female condoms as a way to reduce HIV transmission in the presence of male resistance to male condoms and low female bargaining power. Numerous studies have examined the effects of information interventions which attempt to change preferences or beliefs, or incentive interventions which attempt to change risky sexual behaviour directly (see e.g. Thornton (2008); Dupas (2011); De Walque et al. (2012); Baird et al. (2012); Bjorkman Nyqvist et al. (2015); Duflo et al. (2015)). Many of these studies focus on young women. In contrast, we highlight the importance of considering male preferences in contexts where men typically hold high bargaining power within couples. Medical studies have shown that introducing female condoms alongside male condoms improves protection rates (Fontanet et al., 1998; Vijayakumar et al., 2006; Coman et al., 2013; Mantell et al., 2015), but have largely overlooked the role of intra-household bargaining. Meanwhile Ashraf et al. (2014a) examine the effect of incentives on agents selling female condoms, but do not study impacts on end users.

More broadly, this study contributes to a literature examining the relationship between intra-household bargaining and technology adoption, for example in the form of cookstoves (Miller and Mobarak, 2013; Mohapatra and Simon, 2017), savings accounts (Schaner, 2015), saving through ROSCAs (Anderson and Baland, 2002) and microfinance (Van Tassel, 2004). To our knowledge, we are the first study explicitly to model and estimate the trade-offs inherent in introducing a second-best technology, when low female bargaining power constrains adoption of the first-best technology.

2 Context and intervention

2.1 HIV and condom use in Maputo

Our study took place in Matola, which is the capital of Maputo Province and lies approximately 10km west of Maputo City. HIV prevalence in Maputo Province is high and disproportionately affects women, at an estimated 29.6% for women and 15.8% for men (Ministério da Saúde, 2015). Concurrency among men has been identified as a contributing factor, even among men in stable relationships (Macia et al., 2011). Indeed, 85% of the women in our sample are in stable relationships, but of these 36% report believing their partner is "involved" with other people. In such a climate, technologies which protect against transmission of HIV and other STIs are not close substitutes for pure contraceptive technologies such as the pill, and may be used in addition to pure contraceptive technologies. In our baseline sample, 39% of respondents are currently using pure contraceptive methods (mainly the pill or injectables), and of these 40% are also currently using male condoms.

Both male and female condoms are available in Matola, but male condoms are far more accessible. Female condoms are typically only available at health facilities, which subjects report would take on average 60 minutes to reach, and even there are subject to frequent stock-outs (Pilz, 2014). In contrast, male condoms are readily available, both for free at health facilities and cheaply on the private market. Yet despite the widespread availability of male condoms, there is evidence that men's preferences constrain adoption.[7] Of the women in our study who are currently sexually active but not

[7] As well as being more vulnerable to HIV/AIDS, women also report lower desired current fertility than their male partners: 12% of women in stable relationships say that they want another child now, whereas 23% claim that their partner does. Men may also have higher desired total fertility: 68% of women claim their partner wants another child, whereas only 55% of the women say they themselves want another child.

using any form of protection at baseline, by far the most common reason given is that their partner does not like to or refuses to use them (45% of responses).

2.2 Female condom intervention

Evidence suggests that small-group information and education interventions may be particularly important for promoting female condom use (Terris-Prestholt and Windmeijer, 2016). The intervention we study is run by Pathfinder International, and is aimed at women in populations with high HIV transmission risk. The programme consists of six group sessions lasting ninety minutes each, held fortnightly over a three month period. Pathfinder trains female health workers from the local area to facilitate the programme, and thus facilitators are socially proximal to the participants. The sessions cover: information on female condoms and demonstration of their use on pelvic models; information about other contraceptive methods; information on HIV/AIDS and other STIs; and discussions around consent, negotiation of contraceptive use, intimate partner violence, and women's rights.[8] Group sizes range from a minimum of five to a maximum of twelve women per facilitator, which are thresholds set by the NGO for creating an environment conducive to discussion. Female condoms are also added to the set of products carried by local health workers — which already include male condoms — that participants can access freely and discreetly at the end of each session.

The intervention thus allows us to study which women, if any, adopt female condoms when informational, supply and price constraints are alleviated. The estimated treatment effect may also include the effect of simply coming together in a group with other women to discuss personal issues. We do not attempt to disentangle these mechanisms, since our primary objective is to study how bargaining power affects adoption of female condoms once all constraints to adoption apart from intra-household bargaining are alleviated. Moreover, any standard sex education programme would likely involve all of these components, their combined is arguably of most interest to policymakers.

[8] Qualitative evidence from the medical literature suggests that information about use and about negotiation help introduction of female condoms (Schuyler et al., 2016). The discussions are also included for ethical reasons, to mitigate any risk of these women facing increased violence when introducing new contraceptives into the home.

3 Theoretical framework

In this section we introduce a simple model of intra-household bargaining over STI protection technologies. We abstract from pure contraceptive technologies such as the pill, in light of the evidence discussed above that these are not close substitutes for STI protection methods in contexts such as our study setting. We use the model to formalise two main predictions about the effects of making female condoms freely available: first, on the intensive margin, that female condoms will be adopted by women with intermediate bargaining power; and second, on the extensive margin, that this will lead to an increase in the probability that couples have sex. For ease of representation, we present the model here without the possibility of intra-household transfers – for example, if one partner offers to do more household chores in order to compensate the other partner for a given choice of contraceptive technology. The Online Appendix shows that all of the predictions are robust to generalising the model to allow for transfers, as long as those transfers are not perfectly frictionless: a reasonable assumption if there are utility costs to negotiating transfers, or productivity losses from overriding the usual division of chores within the household.

Preferences: Consider a population of heterosexual couples each consisting of a male m and a female f. When considering the choice of STI protection technology, individual i has preferences over the levels of pleasure and health that the technology yields on average to the population, $u_i(P, H)$, which is quasi-concave and increasing in each argument. For example, P may include the average level of discomfort associated with the material used to produce the technology, and H may include the average level of HIV transmission risk provided by the technology. We allow for idiosyncratic and gender-specific heterogeneity in preferences over P and H through the utility functions. For example, an individual may place a larger weight on health if she is particularly risk-averse, or believes that she has a particularly high risk of HIV infection due to her beliefs about her partner's sexual behaviour. However, we assume that on average, couples' preferences satisfy the following single-crossing property:

Assumption 1.

$$\frac{\partial u_m(P, H)/\partial P}{\partial u_m(P, H)/\partial H} > \frac{\partial u_f(P, H)/\partial P}{\partial u_f(P, H)/\partial H} \tag{1}$$

That is, we argue that the marginal rate of substitution between health and pleasure is greater for women than for men. This assumption is motivated by the facts discussed above, that women on average face greater risk of contracting HIV and greater costs from pregnancy than men do, and that men have stronger reported displeasure and stigma from condom use.

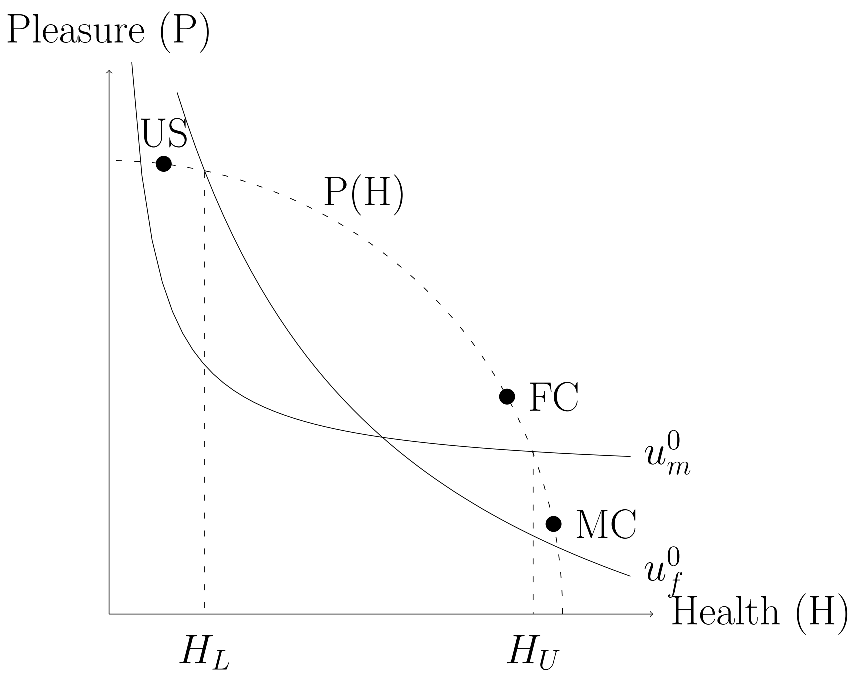
Technologies: In general, let the STI protection technology frontier be represented by a continuously-differentiable function $P(H)$ for $H \in \left[\underline{H}, \overline{H}\right]$. By definition of being on the frontier, $P'(H) < 0$, and let $P''(H) \leq 0$ such that the frontier is weakly concave. This is illustrated in Figure 1. In reality, only certain points on the frontier are easily accessible to couples, depending on the technologies that are readily available.[9] For simplicity, we assume that prior to our intervention, the set of readily available technologies is just the binary set of points on the frontier $\{US, MC\}$. Male condoms (MC) offer greater health than unprotected sex (US) because of their protection against HIV/AIDS and other STIs, but offer lower pleasure.

Our treatment expands the set of readily available technologies to the ternary set of points on the frontier $\{US, FC, MC\}$. As discussed in Section 1, female condoms (FC) provide lower effectiveness and thus lower health than male condoms, but are considered more pleasurable by both genders. For both men and women, they hence represent an intermediate option between male condoms and unprotected sex, as shown in Figure 1. Of course, couples may have initial uncertainty about the pleasure and health associated with female condoms. In what follows we abstract from such uncertainty and consider the permanent adoption decision, once learning has taken place.

Co-operative decision-making: The decision to use an STI protection method must be taken jointly, since use of both male and female condoms is observable. The decision must also be taken each time a couple has sex, unlike decisions over longer-acting contraceptives. In stable couples (who comprise 85% of our sample), the decision to use an STI protection method can thus be thought of as a repeated game with a long

[9]Couples could mix their use of two or more technologies such as to obtain a wider range of points on the frontier. By abstracting from such solutions, we are effectively assuming that due to transaction costs from mixing, the range of points available on the frontier is larger the greater the number of technologies available, and couples prefer to adopt a new technology that yields a given point rather than mixing two other technologies to obtain that point.

Figure 1: STI protection technology frontier



time horizon. It is therefore natural to make the following modelling assumption:

Assumption 2. Decisions over STI protection technologies are taken co-operatively, resulting in choices that are Pareto efficient.

This implies that if a couple decides to have sex, they choose H (and consequently P) to maximise the following household utility function

$$V = \alpha u_f \left(P(H), H\right) + (1 - \alpha)\, u_m \left(P(H), H\right), \tag{2}$$

where $\alpha_i \in [0, 1]$ is the woman's Pareto weight in the couple's sharing rule (Browning and Chiappori, 1998). We interpret this weight as her bargaining power, and thus Equation 2 represents a collective model of the household (Chiappori, 1992). The weight α may depend on factors such as the woman's relative contribution to the couple's income and housework, and her options outside of the relationship. As a simplification, we also assume that the financial and opportunity costs of acquiring any of the technologies is zero. This is true in our experimental setting, and in most public health programmes, where male and female condoms are made available for free if they are provided.

Intensive margin: It is straightforward to show that as long as Assumption 1 holds, the optimal choice of health is increasing in α. The Online Appendix contains the full proof, and proof that this is still the case if partners can make transfers to one another, as long as these transfers are not perfectly frictionless. The intuition is simple: if the female places relatively greater weight on health than the male does, then the more bargaining power she holds, the more the household's choice of STI protection technology will be tilted towards health, and consequently away from pleasure.

Given this result, our first prediction follows straightforwardly (again, full proof in Online Appendix):

Proposition 1. Women with intermediate bargaining power will adopt female condoms.

Intuitively, women with the lowest bargaining power will remain having unprotected sex, and women with the highest bargaining power will continue to use male condoms since they offer the highest health protection.[10] Women with intermediate bargaining power will convince their partners to use female (but not male) condoms.

Note that this does not mean that we will necessarily observe an "inverse-U" relationship between bargaining power and female condom adoption in our experimental sample. The relationship observed in the sample will depend on the extent to which women from the full distribution of bargaining power sign up for the condom programme. Women with the lowest bargaining power may not select into our sample if their partners disapprove of them participating, or indeed if they predict that they will not be able to convince their partners to use female condoms even after the training. If so, then we may not observe the upward-sloping portion of the inverse-U relationship between bargaining power and female condom adoption, but instead observe a simple negative relationship. On the other hand, women with the highest levels of bargaining power may not sign up, if they are already able to persuade their partners to use male condoms. If so, then we may observe a simple positive relationship between bargaining power and female condom adoption.

In terms of the margins of adoption, both couples who were previously having un-

[10] In our specific intervention which educates women only, some women with higher bargaining power who had already persuaded their partners to use male condoms might not even find it incentive-compatible to tell their partners about the availability of female condoms.

protected sex and couples who were previously using male condoms may adopt female condoms, if this interior option allows them to get closer to their optimal point on the technology frontier. Among the women who are engaging in unprotected sex at baseline, women with relatively higher bargaining power — although still relatively low bargaining power compared to the whole distribution — will take up female condoms. Among women using male condoms at baseline, women with relatively lower bargaining power will switch from male to female condoms. Which effect dominates empirically is an important question. If take-up of female condoms mainly comes from women who were engaging in unprotected sex at baseline, then introducing female condoms unambiguously increases rates of protection against HIV/AIDS and other STIs. On the other hand, if female condoms are mainly used as substitutes for male condoms, then offering female condoms will not lead to an increase in condom coverage. Whilst couples who switch to female condoms must be better off in terms of their private utility, the marginal loss of effectiveness is likely to reduce welfare from the perspective of a social planner, given the negative externalities inherent in transmission of HIV and other STIs.

Extensive margin: Let $s \in \{0, 1\}$ indicate the choice of whether to have sex or not. The no-sex option $s = 0$ can be enforced by either partner, and gives reservation utility u_i^0 to each partner. This can be thought of as the utility from partners' best immediate alternative, for example in terms of time use. Along with $s = 1$ is a choice of contraception from the available sets as described above.

It is straightforward to see that the introduction of female condoms increases the probability that both couples' reservation utilities are satisifed and hence that $s = 1$; see Online Appendix for formal proof. This leads to our second prediction:

Proposition 2. Making female condoms freely available increases the probability that couples have sex.

To illustrate, Figure 1 depicts a couple whose reservation utilities are only both satisfied following the introduction of female condoms.

Note that bargaining power α does not enter a couple's decision as to whether to have sex or not: this extensive-margin decision depends only on individual reservation utilities and preferences over pleasure and health, and the set of readily available points on the technology frontier. Thus the effect of the intervention on the extensive margin decision is not predicted to vary by female bargaining power, conditional on these factors.

4 Study design and data sources

4.1 Study design

Pathfinder International began its female condom programme in Matola in 2011. We expanded the progamme to four additional neighbourhoods in 2014, using a phased-in experimental design. The healthcare workers who facilitated the programme first conducted door-to-door recruitment to identify women willing to participate. The eligibility criteria were that women needed to be between 18 and 49 years of age, sexually active, and not pregnant. To avoid spillovers, if a woman signed up with a friend or relative then they were both included in the programme but were automatically assigned to a separate set of groups. These groups did not form part of the study, but still received the intervention at a later stage. Spillovers between unconnected individuals are unlikely to be a cause for concern, since the average population of these neighbourhoods is 20,000 inhabitants. Our final baseline sample consists of 298 respondents, who received the baseline survey in August 2014.[11]

At the end of the baseline survey, it was explained to each participant that assignment to the first or second wave of the programme would be determined by computer randomisation. Once the entire sample had responded to the baseline survey, the research team randomly allocated half of the respondents recruited by each facilitator to the treatment group and half to the control group.[12] The reason for stratifying on facilitator was to improve power, and to ensure that there would be enough space for treatment and control participants to attend sessions close to their home.

The treatment group then received the intervention from September-December 2014. The endline survey was conducted in February-March 2015, five to six months after the

[11]Overall 317 women were recruited into the study. However, one facilitator fell severely ill shortly after the baseline, and there was nobody sufficiently trained to replace her. Thus all 35 individuals recruited by her (treatment and control) were dropped from the study.

[12]The randomisation was done in private, given the sensitive nature of participating in our intervention. A member of the research team took the list of respondents for each facilitator, sorted them by a randomly-generated number, and assigned the first half to treatment and the second half to control.

intervention had started for treated individuals, and two to three months after treated individuals had received their last group session. A total of 232 respondents were traced and administered the endline survey. The retention rate was thus 78%, which is similar to that in other studies tracking female populations in urban or peri-urban areas (see e.g. Banerjee et al. (2015); Cohen et al. (2017)). The Online Appendix shows that the few observable predictors of attrition are not differential across treatment and control groups. However, Table 1 shows that attrition is higher in the control group (27% in the control group versus 18% in the treatment group, significant at the 10% level). Section 5.4 details a number of robustness tests that we perform to account for this.

Following the endline survey, the control group then received the intervention from March-May 2015. Due to administrative complications in rolling out the endline survey, the control group for one facilitator was administered the endline survey after having been treated. These five observations are dropped from all estimations of treatment effects, leaving a final estimating sample of 227 respondents. All effects estimated below are "intent-to-treat", given that attendance was far from perfect among those respondents assigned to treatment.

4.2 Survey data

Table 1 shows measures of key covariates and contraceptive use elicited for the full baseline sample.[13] All are balanced across treatment and control. 85% of respondents report being in a stable relationship with average duration of 8.7 years, comprising 63% are married and 22% who are unmarried but still in relationships of on average 4.8 years.[14] The rest of the sample (15%) are sexually active but not in a stable relationship. The vast majority of respondents report having had just one sexual partner in the last twelve months, with 10% reporting zero partners and 3% reporting two partners. A third of respondents report being HIV-positive, which is close to the official statistics reported above. Slightly more than 10% of respondents report having had an STI in

[13]Baseline covariates are also balanced when attritters are excluded, see Online Appendix.

[14]The latter includes traditional marriages and respondents who describe themselves as "living as married" but not legally married. The latter is common in this region due to the high bride price and costs of obtaining a marriage certificate.

the last month; although this may be under-reported.[15] Fewer than half, 41%, are able to name the female condom when asked to list contraceptive methods.

Our primary outcome variables are the use of contraceptive methods, disaggregated by female condoms, male condoms and other modern contraceptive methods (mainly the pill and injectables). For each method, we ask respondents whether they have ever used that method, and whether they are currently using it (i.e. consider it to be part of their current portfolio). For male and female condoms, we also ask whether they have used that method in the last thirty days. Table 1 describes the baseline values of each of these measures. Baseline use of female condoms is low: 9% of the respondents have ever used a female condom, 3% have used one in the last 30 days, and 2% are currently using female condoms. Male condom use is substantially higher: around three quarters of women have ever used a male condom, 32% have used one in the last 30 days, and 39% percent say they are currently using male condoms. Altogether, 39% of our sample are currently using pure contraception methods at baseline, comprising 20% using the pill and 14% using injectables, and a small number using intrauterine devices (IUDs), the diaphragm, and sterilisation. These women may have signed up to the female condom programme as a way to switch out of these methods, or because they are seeking additional protection against HIV/AIDS and other STIs.

Finally, Table 2 compares our sample to a representative urban sample of women from Maputo Province, from the 2011 Demographic Health Survey data (DHS, 2011). Our sample is close to the overall adult female population of Maputo Province in terms of demographic characteristics such as age, years of education, marital status, pregnancy, and desired fertility.[16] On the other hand, the women in our study appear to have greater bargaining power than the representative sample: they began to have sex at a later age, are more likely to have used a condom the last time they had sex, and report

[15] We do not test for HIV, since we would have low power to measure an impact over the length of our study: official guidelines require a minimum of four weeks after exposure before the first test, and a second test three months later in the case of a negative result for high-risk cases (https://www.gov.uk/government/publications/time-period-for-hiv-testing-position-statement). We also opted not to test for STIs such as chlamydia, given the already sensitive nature of participation in the study and the budgetary implications of providing treatment to those who test positive (as required by medical research ethics guidelines).

[16] One exception is that the women in our sample are much less likely to have a job, which makes sense if women with a lower opportunity cost of time are more willing to participate in a time-intensive programme.

Table 1: Baseline balance of covariates and use – full sample

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age	30.32	30.12	30.52	-0.42	298	146	152
Years of education	6.21	6.26	6.17	0.27	295	146	149
Literate	0.84	0.84	0.85	-0.17	295	144	151
HH head	0.22	0.21	0.24	-0.51	298	146	152
Income							
Has job	0.38	0.42	0.33	1.64	295	144	151
Personal income last 30 days	745.85	854.52	641.46	1.41	298	146	152
Relationships							
In a stable relationship (incl. married)	0.85	0.85	0.84	0.17	298	146	152
Married (officially or unofficially)	0.63	0.64	0.62	0.37	297	146	151
Years relation	8.66	8.62	8.70	-0.08	235	114	121
# Partners last 12 months	0.92	0.92	0.93	-0.23	298	146	152
Sexual behaviour							
Pregnant	0.05	0.05	0.06	-0.42	297	145	152
HIV positive (self-report)	0.33	0.35	0.31	0.75	260	129	131
STI last 3 months (self-report)	0.13	0.13	0.13	-0.10	259	124	135
Names FC as contraceptive	0.41	0.44	0.39	0.90	296	146	150
Baseline use							
Ever used FC	0.09	0.09	0.09	0.11	298	146	152
Ever used MC	0.74	0.76	0.73	0.59	298	146	152
Ever used other	0.72	0.72	0.72	0.04	298	146	152
Used FC last 30 days	0.03	0.01	0.04	-1.39	298	146	152
Used MC last 30 days	0.32	0.28	0.35	-1.26	298	146	152
Current use FC	0.02	0.02	0.03	-0.33	298	146	152
Current use MC	0.39	0.37	0.41	-0.79	298	146	152
Current use other	0.39	0.41	0.37	0.75	298	146	152
Attrition							
Interviewed at endline	0.78	0.73	0.82	-1.86*	298	146	152

Notes: N=298 in the baseline sample. Lower sample sizes reflect observations that are missing or not applicable. "Treatment Mean" contains all individuals assigned to the treatment group, whether or not they attended the sessions. Column 5 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.

greater decision-making power. This is in line with the suggestion in Section 3 that women with very low bargaining power might not select into our study.

Table 2: Comparison of Study Sample to DHS Representative Sample

	Study Mean	DHS Mean	t-test	Study N	DHS N
Demographics					
Age	30.18	29.47	1.28	276	1007
Years of education	6.35	6.72	-1.82	273	1007
Literate	0.85	0.76	3.82	273	1007
Income					
Has job	0.37	0.58	-6.33	273	1007
Relationships					
Married (officially or unofficially)	0.63	0.61	0.58	275	871
Pregnant	0.05	0.07	-0.76	275	1007
Wants another child in future	0.57	0.57	0.14	260	961
Sexual Behaviour					
Age of sexual debut	16.62	16.16	3.26	273	955
Used condom during last time sex	0.54	0.31	6.32	243	871
Decision-making visiting family	0.62	0.39	6.43	272	580
Decision-making spending earnings	0.59	0.21	10.85	275	569
Decision-making her health	0.53	0.39	4.04	275	580

Notes: Column 2 displays the mean from our study, N=298. Column 3 shows the 2011 DHS mean for women in urban areas of Maputo Province, N=1007. Lower sample sizes reflect observations that are missing or not applicable. Variables selected for comparison are those that appear in both our study and the DHS, with similar or identical wording.

4.3 Diary data

A voluntary subsample of the survey respondents also took part in weekly coital diaries. Altogether 56 respondents volunteered to participate, comprising 27 in the treatment group and 29 in the control group (see Online Appendix for tables of balance on covariates and baseline use within this sample).[17] The diaries recorded detailed information on all of the respondents' sexual encounters in the seven days prior to each interview.

[17]The Online Appendix also shows that the diary participants are representative of the balanced panel of all survey participants, with a few exceptions: the diary participants have been in a relationship for longer than the average study participant, and no diary respondents are pregnant. The results from the diary subsample presented below are robust to re-weighting to make the diary subsample representative of the full sample (available on request).

Diary interviews took place with carefully trained enumerators in a private place chosen by the respondent. The same enumerator interviewed a given respondent each week, to maximise trust and confidentiality. Diary interviews took place over a period of 17 weeks, beginning four weeks prior to the first group receiving its first session and ending one week after the last group received its last session. The baseline period for each respondent is taken to run from the start of the diary data collection until one week after the facilitator that the respondent was assigned to began her first meeting for her treatment-group participants (5.6 weeks on average). The endline period is taken to run from the end of the baseline period until the end of the diary data collection, comprising 8.9 weeks on average. On average 75% of the diary sample participated each week.[18]

Although the diary data come from a small and potentially selected sample, they allow us to provide illustrative evidence on a number of important concerns. First, the fact that the diaries are administered very shortly after the sex acts that they ask about reduces recall bias (Das et al., 2012). The diaries are also a more complex and granular instrument than the baseline and endline surveys. Although the time period covered by the surveys was slightly different to that covered by the diary interviews, we are able to cross-check with the surveys to reduce concerns about misreporting in either instrument. There is limited evidence of under-reporting of contraceptive use in the surveys: 5 out of 56 diary participants report never having used a female condom during the endline survey but report using them in the diaries; whilst for male condoms the figure is 4 out of 56 respondents.[19] We therefore run the main impact analyses using both the survey data and the diary data; see Section 5 for details.

As well as constructing variables for contraceptive use at the respondent level, the diary data allow us to analyse the impact of the intervention at the level of the sex act. Altogether respondents report a total of 349 sex acts during the endline period: an average of 6.1 sex acts per respondent, with a minimum of zero and a maximum of 30.

[18]Individual respondents took part in the diaries an average of 13 times, with a minimum of three weeks and a maximum of 17 weeks. There are no significant differences in participation between the treatment and control group.

[19]We cannot make the opposite comparison, given that the endline survey took place two months after the end of the diaries: if a respondent reports using condoms in the survey but not the diaries, it may be that she adopted them during those two months.

4.4 Bargaining power

To test the model's predictions, we require proxies of women's bargaining power within a relationship. By virtue of their definition, we elicit these measures only for women with stable partners, although as mentioned earlier this comprises 85% of the sample. We first include a survey module about assets brought by the woman to the relationship. This has the advantage of proxying bargaining power in a way that is unconfounded by current contraceptive use. We next include a standard survey module on decision-making, and a novel survey module on power dynamics in the relationship, which we developed through extensive piloting in the local context. These modules have the advantage of proxying bargaining power more directly and more currently, although such elements of bargaining may be simultaneously determined along with contraceptive use. Table 3 provides summary statistics for each of the questions in each module.

Since each of these modules contains multiple questions that are highly correlated, we first perform a principal component analysis to construct indices. The full tables of component loadings can be found in the Online Appendix. The first three principal components of the assets module jointly explain 68.5% of the variance in the assets questions. Similarly, we run a principal component analysis including the questions from the decision-making and power dynamics modules. These two survey modules each load a different principal component, which together explain approximately 40% of the variance in the survey questions. Thus altogether we take five principal components as our preferred measures of bargaining power: three from the assets module, and one each from the decision-making and power dynamics modules. For ease of comparability, we scale the components so that the woman with least bargaining power on that measure has a score of zero, and normalise them such that a one point increase in each measure represents an increase of one standard deviation. The Online Appendix shows that these principal components are balanced across treatment and control.

Reassuringly, these measures of bargaining power are correlated with baseline demographic characteristics in the way that one would expect: both the second and the third principal components of assets brought by the respondent to the relationship are strongly correlated with her education; the respondent's decision-making power is strongly positively correlated with her personal income in the last thirty days, age, whether she is the household head, and whether she has a job; whilst her decision-

Table 3: Bargaining power – summary statistics

	Mean	sd	Min	Max	Total
Did you bring the following assets to your relationship...					
...jewellery?	0.08	0.27	0.00	1.00	264
...animals?	0.00	0.06	0.00	1.00	264
...land?	0.01	0.11	0.00	1.00	264
...electronics?	0.02	0.14	0.00	1.00	264
...money?	0.07	0.25	0.00	1.00	264
...mobile phone?	0.07	0.25	0.00	1.00	264
...kitchen utensils?	0.10	0.30	0.00	1.00	263
Who decides about...					
...buying clothes for you?	0.80	0.40	0.00	1.00	297
...buying phone credit?	0.76	0.43	0.00	1.00	297
...education for the children?	0.49	0.50	0.00	1.00	288
...health expenses for you?	0.55	0.50	0.00	1.00	297
...health expenses for the children?	0.41	0.49	0.00	1.00	291
...if you are allowed to work?	0.59	0.49	0.00	1.00	296
...how earnings are used?	0.60	0.49	0.00	1.00	297
...visits to friends?	0.64	0.48	0.00	1.00	296
...visits to family?	0.64	0.48	0.00	1.00	294
Who usually has more say when you talk about serious things	0.47	0.50	0.00	1.00	250
In general, who do you think has more power in your relationship	0.39	0.49	0.00	1.00	249
Power dynamics					
Most of the time, we do what my partner wants to do	2.33	1.08	1.00	4.00	250
My partner won't let me wear certain things	2.61	1.11	1.00	4.00	250
When my partner and I are together, I'm pretty quiet	3.07	0.96	1.00	4.00	250
My partner has more say about important decisions that affect us	2.39	1.09	1.00	4.00	250
My partner tells me who I can spend time with	2.79	1.09	1.00	4.00	249
I feel trapped or stuck in our relationship	3.20	0.86	1.00	4.00	250
My partner does what he wants, even if I do not want him to	2.86	1.00	1.00	4.00	249
I am more committed to our relationship than my partner is	2.74	1.08	1.00	4.00	250
My partner is involved with other people apart from me	2.77	1.02	1.00	4.00	249
My partner always wants to know where I am	2.16	1.10	1.00	4.00	250
When my partner and I disagree, he gets his way most of the time	2.73	1.06	1.00	4.00	248

Notes: All values taken from the baseline survey (N=298). All variables are coded such that a higher value proxies greater bargaining power for the respondent. The assets module 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). The decision-making module was enumerated to all respondents (N=298), except the questions "who has more influence" 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). Lower observation numbers reflect values missing or unwillingness to answer.

making power is strongly negatively correlated with whether the couple is married. One anomaly is that her decision-making power is negatively correlated with her education. To avoid the measures of bargaining power spuriously proxying the effects of any of these demographic characteristics, we therefore add these characteristics as controls when estimating the effects of bargaining power on condom adoption; see Section 5.3.

As predicted by the model, the measures of bargaining power are strongly negatively correlated with the respondent reporting that her partner refuses to use male condoms. Specifically, the correlation between "partner refuses male condoms =1" and a one standard deviation increase in the principal component of bargaining power in the following domains are: power dynamics -0.11 (p-value 0.09); assets 1 -0.13 (p-value 0.03); assets 2 -0.10 (p-value 0.002); assets 3 -0.18 (p-value <0.01). We also observe a positive correlation between assets brought by the woman to the relationship and baseline use of male condoms in the last thirty days: a 7.3 percentage point increase in the likelihood of use per one standard deviation in assets 1 (p-value 0.035).

5 Results

5.1 Impacts on condom use

Our preferred estimations are analysis of covariance (ANCOVA) linear probability models of the following form:[20]

$$Pr\left[Y_{if1} = 1 | Y_{if0}, treat_{if}, \eta_f\right] = \alpha + \delta Y_{if0} + \beta treat_{if} + \eta_f, \tag{3}$$

where Y_{if1} is the outcome variable of interest at endline, and Y_{if0} is its value at baseline. $treat_{if}$ is a dummy for being assigned to the treatment group, i.e. to receiving the programme in the first rather than the second phase. β represents the intent-to-treat effect, since not all individuals assigned to treatment attended the programme: the participation rate was around 65% for each individual session, with 20 women (17.7% of the treatment group) not attending any of the six sessions. 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-

[20]Results are robust to using logit, Firthlogit or Relogit specifications, see Online Appendix.

level heteroskedasticity, as this was the level of randomisation (Abadie et al., 2017). Given concerns about the use of asymptotic results about standard errors in smaller samples, we also report additional p-values for the treatment coefficients as calculated from randomisation inference tests (Young, 2016).

Table 4 displays the main impacts of the intervention. The programme has a substantial and highly significant effect on the use of female condoms: we observe an 18.4 percentage point increase in the proportion of women who have ever used female condoms (equivalent to 209% of the endline mean in the control group); a 4.7 percentage point (470%) increase in the proportion who have used a female condom in the last thirty days; and a 7.7 percentage point (385%) increase in the proportion who are currently using female condoms. The fact that the treatment effect on ever use is higher than the treatment effect on use in the last thirty days and current use suggests that many women in the treatment group try female condoms at the start of the intervention, then a smaller although sizeable fraction continue to use them. This is a natural adoption pattern if couples experiment with female condoms and thereby learn more about their costs and benefits, then some return to their original contraceptive method while others adopt female condoms more permanently.

We see no evidence of anticipation effects or contamination in the control group, as the control group means of female condom use at endline are still very low and identical to those at baseline (see Table 1). This is unsurprising, since female condoms are difficult to obtain in the study area outside of our intervention. Indeed, the number of free female condoms that a respondent in the treatment group took from the sessions is highly correlated with her report of ever use (correlation 0.38, p-value < 0.01), use last 30 days (0.21, p-value 0.02), and current use (0.29, p-value <0.01). This weighs against concerns that reported use of female condoms might represent response bias.

We do not observe any significant evidence that respondents substitute away from or increase their use of male condoms or other, pure contraceptives, although we are less powered to detect changes in their use.[21] Table 5 shows that when we split the sample by women who are using or not using male condoms at baseline, in line with the

[21] We have 80% power to detect the following minimum detectable effect sizes at the 5% level in a two-tailed test: ever use – female condoms 7.6 pp, male condoms 9.6 pp, other 10.3 pp; use last 30 days – female condoms 3.5 pp, male condoms 14.0 pp; current use – female condoms 4.9 pp, male condoms 13.9 pp, other 13.6 pp.

Table 4: Treatment effects – primary outcome variables

	(1) **Ever FC**	(2) **Ever MC**	(3) **Ever other**	(4) **Last 30 days FC**	(5) **Last 30 days MC**	(6) **Current FC**	(7) **Current MC**	(8) **Current other**
Treatment	0.184***	-0.012	0.020	0.047**	-0.052	0.077**	0.060	0.030
(s.e.)	(0.042)	(0.041)	(0.042)	(0.023)	(0.057)	(0.030)	(0.058)	(0.053)
[r.i. p-value]	[0.000]	[0.777]	[0.649]	[0.080]	[0.359]	[0.025]	[0.348]	[0.583]
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to 8 to whether she is currently using it. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator fixed effects (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 are estimated from Monte Carlo simulations re-assigning treatment within facilitator strata, with 1000 repetitions.

model we see take-up from both groups. The strong take-up of female condoms among women not using male condoms at baseline (columns 1, 3 and 5) is particularly important from a policy perspective. It implies that the treatment decreases the number of women having sex unprotected from HIV/AIDS and other STIs, rather than generating pure substitution away from other STI contraceptive methods. We do not observe any significant impacts on male condom use when we split the sample by those using and not using male condoms at baseline (see Online Appendix).

Table 4 shows that we also see no increase in or substitution away from other contraceptive methods such as the pill and injectables. This suggests that women who adopt female condoms were either previously using no contraceptives, or use female condoms in addition to other methods in order to protect against HIV/AIDS and other STIs. Indeed, of the women who are currently using female condoms at endline, 42% are also using other contraceptive methods (mainly the pill or injectables).

5.2 Extensive-margin impacts

We use the diaries to examine the effects on the extensive margin, i.e. the probability of having sex. Our preferred measure of this is the likelihood of at least one sex act per respondent per week, so that results are not unduly influenced by a very small number of respondents who report a large number of sex acts. Taking advantage of the weekly nature of the diaries, we estimate the following fixed effects panel specification:

$$\begin{aligned} &Pr\left[Y_{ift} = 1 | treat_{if}, \eta_f, \phi_{if}\right] \\ &\quad = \alpha + \delta \times endline_t + \beta treat_{if} \times endline_t + \eta_f \times endline_t + \phi_{if}, \qquad t = 1, 2, ..T \end{aligned} \tag{4}$$

where Y_{ift} is the outcome variable of interest for individual i assigned to facilitator f in week t. The unit of observation is thus the respondent-week. Standard errors are again clustered at the individual level. As discussed earlier, acknowledging the small and potentially selected nature of the diaries sample, we present the following results as illustrative only.

Table 6 shows that, in line with Proposition 2 of the model, the introduction of female condoms leads to a significant increase in the likelihood of sex acts. In the full endline

Table 5: Treatment effects on female condom use, by baseline male condom use

	(1) **Ever FC No MC at baseline**	(2) **Ever FC Current MC at baseline**	(3) **Last 30 days FC No MC at baseline**	(4) **Last 30 days FC Current MC at baseline**	(5) **Current FC No MC at baseline**	(6) **Current FC Current MC at baseline**
Treatment	0.169***	0.232***	0.073**	0.030	0.085***	0.049
(s.e.)	(0.047)	(0.074)	(0.030)	(0.034)	(0.031)	(0.057)
[r.i. p-value]	[0.004]	[0.006]	[0.023]	[0.532]	[0.035]	[0.490]
Observations	141	86	141	86	141	86
Control mean endline	0.092	0.081	0.000	0.027	0.000	0.054

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms. Columns 1-2 refer to whether the respondent has ever used that method, columns 3-4 to whether she has used it in the last 30 days, and columns 5-6 to whether she is currently using it. Odd-numbered columns present results for the subsample of individuals not currently using male condoms at baseline; even-numbered columns present results for the subsample of individuals who are currently using male condoms method at baseline. The bottom two rows present chi-squared statistics and their p-values for the test that the treatment effect is the same across the two subsamples. These are obtained from seemingly unrelated estimations. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator fixed effects (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 are estimated from Monte Carlo simulations re-assigning treatment within facilitator strata, with 1000 repetitions.

period, respondents in the treatment group were on average 9.1 percentage points more likely to report a sex act in a given week, compared to a control group mean of 46.9%. In the last 30 and 14 days, the treatment effect on the likelihood of sex acts per week was 11.3 pp and 15.8 pp respectively, compared to 47.1% and 49.1% in the control group. The fact that we observe this increase in the treatment group indicates that there are couples in which one or both partners' participation constraints are sometimes or always binding when the only options are male condoms or unprotected sex, but where both find sex with female condoms preferable to not having sex. The introduction of female condoms therefore increases utility for such couples. Moreover, the Online Appendix shows that the increase in sex acts is driven by those respondents who are not using male condoms when they do have sex at baseline.

The Online Appendix also shows that we observe a large and highly significant reduction in the proportion of sex acts in which a discussion or disagreement about condoms takes place. This supports the idea that the expansion from a binary to a ternary choice allows the couple to choose an STI protection technology that is closer to their preferred choice on the technological frontier. Reassuringly, in the survey data we see no negative impact of treatment on measures of women's self-reported well-being, nor do we see any impacts on emotional or physical violence (see Online Appendix). This contrasts Ashraf et al. (2014b)'s finding of a negative effect of concealable contraceptives on women's self-reported well-being. The difference is likely attributable to the fact that our treatment expands the set of observable contraceptives available in a cooperative setting; rather than introducing a concealable contraceptive in a non-cooperative setting, which women may choose to use but with some psychic cost of hiding.

We would expect women who are not in a stable relationship to place a larger weight on the health offered by STI protection technologies, and so to have a higher demand for condoms. Indeed, the Online Appendix shows that the treatment effect on ever use of female condoms is stronger for women who are not in a stable relationship at baseline. Nonetheless, when we restrict the sample to just those women in a stable relationship, we still observe positive treatment effects on female condom use: a 16.4 percentage point increase in ever use of female condoms (p-value < 0.01), a 5.6 percentage point increase in use in the last 30 days (p-value 0.042), and a 7.9 percentage point increase in current use (p-value 0.019). This may be rational if one partner is HIV-positive while the other

Table 6: Impacts on likelihood of sex acts per respondent week – 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	0.091**	0.113**	0.158*
	(0.045)	(0.057)	(0.086)
Treat × Facilitator f.e.'s	✓	✓	✓
Observations	863	536	367
Control mean	0.469	0.471	0.491

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for whether a respondent had a sex act in a particular week. "Treat × endline" is an indicator for observations in the treatment group during the endline period. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions are respondent-level LPM fixed effects models with the respondent-week as the unit of observation. All specifications include individual fixed effects (N=56), and facilitator × endline fixed effects (N=17) for inference since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

is HIV-negative, or if one or both partners have relations with others or suspect that their partner does.

5.3 Heterogeneity by bargaining power

We now test our main predictions about which women, among those in stable relationships, adopt female condoms in terms of their bargaining power. Recall that Proposition 1 of the model is an inverse-U relationship between bargaining power and adoption of female condoms in the population. However, given that women in our sample report lower rates of violence and higher rates of decision-making power than the DHS average (see table in Online Appendix), this is strongly suggestive that women with the lowest bargaining power in the population do not select into the condom programme. We would therefore expect the relationship between bargaining power and female condom adoption in our sample to be negative, as we should predominantly capture the downward-sloping portion of the inverse-U relationship for the whole population.[22]

Table 7 displays the relationship between endline current use of female condoms (our preferred proxy of sustained adoption) and each of the baseline principal component measures of bargaining power, both alone and interacted with treatment.[23] We control for baseline female condom use and for the demographic factors which are significantly correlated with baseline bargaining power: age, education, income in the last thirty days, and whether the respondent is the household head. The results are striking: for almost every measure of bargaining power, we observe a negative interaction between baseline

[22]We have power to detect reasonably modest increases in female condom use among the least empowered women in our sample — corresponding to women with intermediate levels of power in the population as per Proposition 1, as explained above — as well as among the most empowered women. If we split the sample along each measure of bargaining power, we have 80% power to detect the following minimum detectable effect sizes at the 5% level in a two-tailed test when the dependent variable is current use of female condoms and the same controls are included as in Table 7: assets 1 – below or equal to median 7.2pp, above median 0.0pp; assets 2 – below or equal to median 7.0pp, above median 0.0pp; assets 3 – below or equal to median 6.8pp, above median 0.0pp; decision-making – bottom quartile 13.3pp, top three quartiles 4.7pp; power dynamics – bottom quartile 0pp, top three quartiles 6.7pp. Since each of the assets indices is heavily right-skewed, a median split corresponds to a split by those who have a value of zero or greater than zero. Decision-making and power dynamics are much less skewed; hence we report power calculations for the bottom quartile in order to isolate the women with the lowest bargaining power in our sample.

[23]The seventeen facilitator fixed effects are no longer included in Table 7, due to low number of observations per facilitator to examine heterogeneity along the distribution of bargaining power.

bargaining power and treatment, showing that women with the lowest bargaining power in our sample are the most likely to adopt female condoms as a result of the intervention. As a robustness check, we also show that a post-double LASSO specification in which we include all survey variables and their interaction with treatment selects four out of the five bargaining power measures and their interaction with treatment, as well as age, education and their interactions with treatment (see Online Appendix).

In terms of margins of adoption, as predicted by our model we observe a strong negative correlation between baseline bargaining power and female condom adoption *conditional on using male condoms at baseline* (see Online Appendix). On the other hand, as discussed earlier, we do not see evidence of a large degree of substitution away from male condoms overall. A possible explanation is that women with higher bargaining power who take up female condoms also intersperse their use with the use of male condoms. Indeed, 81% of women who are currently using female condoms at endline also report currently using male condoms. This "double protection" is a typical pattern of adoption observed in the medical literature, and is found to be associated with a large increase in the number of protected sex acts (Vijayakumar et al., 2006).

5.4 Alternative explanations and robustness

Access: A possible alternative explanation for the negative interaction terms could be if women with lower bargaining power are less able than women with higher bargaining power to access male condoms (or other contraceptives) through the market or at health clinics. However, if this was the case then we would expect to see a similar pattern of heterogeneity in current use of male condoms, which the health workers also carry. Yet this is not the case: women with lower bargaining power are not consistently more likely than women with higher bargaining power to take up male condoms (see table Online Appendix). Furthermore, we also re-run estimations of the bargaining power heterogeneity controlling for walking distance to the nearest health centre — a proxy of pre-treatment access to male condoms — and its interaction with treatment. These terms are not significant for female or male condom adoption, and do not significantly change the coefficient on the interaction of bargaining power and treatment.

Table 7: Impacts on current use of female condoms – heterogeneity by bargaining power

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
Treatment	0.128***	0.111***	0.159***	0.223**	0.193*
	(0.045)	(0.039)	(0.051)	(0.093)	(0.114)
Assets 1	-0.002				
	(0.010)				
Treat*Assets1	-0.044**				
	(0.017)				
Assets 2		0.004			
		(0.010)			
Treat*Assets2		-0.031**			
		(0.015)			
Assets 3			-0.001		
			(0.005)		
Treat*Assets3			-0.054***		
			(0.018)		
Decision-making				-0.015	
				(0.022)	
Treat*Decision				-0.079**	
				(0.038)	
Power dynamics					0.014
					(0.019)
Treat*Power dynamics					-0.037
					(0.039)
Controls	✓	✓	✓	✓	✓
Observations	198	198	198	180	180
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Excluding attritters, N=206 women are in a stable relationship at baseline. N=198 are women who are in a stable relationship and have no missing values on the control variables. N=180 have non-missing values for all of the decision-making, power dynamics, and control variables. "Treatment" is a dummy for assignment to treatment 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 identified in Table A.5. "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 A.6. Dependent variables are binary indicators for current use of female condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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.

Use of other contraceptive methods: The interaction between bargaining power and treatment is also not proxying a differential effect of treatment depending on whether the respondent is using other methods of contraception (i.e. the pill or injectables) at baseline. Our measures of bargaining power are positively correlated with current use of the pill at baseline, and negatively correlated with use of injectables at baseline, consistent with the idea that less empowered women use injectables instead of the pill as injectables are more concealable Ashraf et al. (2014b). However, when baseline use of other forms of contraception and its interaction with treatment is included into the regressions, the interactions between treatment and bargaining power remain negative and highly significant (see Online Appendix).

HIV status: Heterogeneity by bargaining power is also not proxying the observed heterogeneity by HIV status. This could have been the case since women with lower bargaining power are more likely to be HIV-positive. However, the interaction of the bargaining power measures with treatment remain negative and significant when controlling for HIV status and its interaction with treatment (see Online Appendix).

Risk beliefs: Finally, heterogeneity by bargaining power is not proxying heterogeneity by women's beliefs about the risk of contracting HIV. This could have been the case if women with lower bargaining power revised their risk perceptions upwards more strongly as a result of the intervention. However, again the interaction of the bargaining power measures with treatment remain negative and significant when controlling for baseline risk beliefs and their interaction with treatment (see table in Online Appendix).

Attrition: As mentioned earlier, attrition is higher in the control group. To account for this, Table A.1 in Appendix A shows the main estimation results with observations re-weighted by the inverse of the predicted probability of an individual appearing in the endline sample, based on her treatment status and her baseline covariates. The results are robust to this re-weighting. Table A.2 Appendix A also presents Lee Bounds for the main impact estimations (Lee, 2009). Note that we do not condition on facilitator fixed effects, since whether attrition is higher in the treatment or control group varies by facilitator, and thus the monotonicity assumption required for conditioning fails. This means that the bounds presented are particularly conservative. Whilst the lower bounds cannot rule out a treatment effect of zero for the various measures of female condom use, we are able to rule out any sizeable negative effects, and the upper bounds are large and highly significant. We also note that the lower bounds are likely to be overly pessimistic in the context of our study, since it was almost impossible for those who attritted from the control group to obtain access to female condoms.

6 Cost-Benefit and Cost-Effectiveness Analysis

To understand how our results might combine to impact welfare and policy, it is important to weigh the increase in condom coverage — and associated reduction in negative externalities from HIV transmission — against the decrease in average condom effectiveness compared to pure use of male condoms, and the observed increase in the number of sex acts. As an illustrative exercise, we conduct a cost-benefit analysis of two possible scale-ups to the entire female population of South Mozambique: a scale-up of our full intervention; and a scale-up of just the free distribution of female condoms, with the assumption that information about female condoms can be provided with zero marginal cost via existing sex education programmes. The purpose of this exercise is to highlight the potential magnitudes of the trade-offs involved in introducing a second-best technology, and the quantitative importance of the behavioural response. The purpose is not to provide an accurate cost-benefit assessment, given the inherent uncertainty in extrapolating from our observed treatment effects to what treatment effects would be in the whole population, over a longer time horizon, and from a different version of the intervention in the case of provision via existing sex education programmes.

The Online Appendix details the methodology of our cost-benefit analysis in full. We adjust the epidemiological model used by UNAIDS in order to estimate the number of HIV infections and disability-adjusted life years (DALYs) that free access to female condoms would help avert, based on our observed treatment effects. We also factor in productivity gains from a reduction in the burden of HIV, as is standard in the literature. On the cost side, we consider programme costs of introducing female condoms, but also cost savings from reduced provision of anti-retroviral therapies and prevention of mother-to-child transmission treatments.

The results show that accounting for the behavioural response, i.e. the observed increase in the number of sex acts, is crucial. Before accounting for this, both our full programme and adding female condoms to existing sex education programmes actually imply a cost saving. Intuitively, this is because low female bargaining power implies that the main margin of female condom adoption is from women previously having unprotected sex, rather than substitution away from male condoms. However, once we incorporate the behavioural response, only adding female condoms to existing sex education programmes has the potential to be cost-effective in our illustrative simulations.

7 Conclusion

Our results suggest that low female bargaining power indeed acts as a key constraint to adoption of male condoms in contexts such as Mozambique. If female condoms are introduced with adequate information and support, they are taken up by women with lower bargaining power who are otherwise having unprotected sex. In terms of policy, this means that the correct cost comparison for free provision of female condoms is not to the free provision of male condoms, but rather to the costs of ARTs and other costs associated with unprotected sex. However, evidence from a similar intervention with a representative sample of the population and a longer time horizon after adoption are crucial to aid more comprehensive cost-benefit calculations and funding decisions.

More broadly, we have highlighted how low female bargaining power may constrain adoption of potentially welfare-improving household technologies, in cases where women have a stronger preference for adoption or face higher costs of non-adoption compared to men. There are many other examples of technologies where women may have a stronger willingness than men to adopt. For instance, women may have a higher demand for insurance, given evidence that they are more risk-averse. In such cases, information and social norm campaigns targeted specifically at men may be the first-best approach to increasing investments and adoption. Otherwise, providing inferior versions of the technology that are more acceptable to men, or bundling technologies with goods for which men have strong demand, may offer a second-best solution. These remain important topics for future research.

References

A. Abadie, S. Athey, G. W. Imbens, and J. M. Wooldridge. When should you adjust standard errors for clustering? *NBER Working Paper No. 24003*, 2017.

J. R. Agnew, L. R. Anderson, J. R. Gerlach, and L. R. Szykman. Who chooses annuities? An experimental investigation of the role of gender, framing, and defaults. *American Economic Review*, 98(2):418–422, 2008.

T. Amemiya. Some theorems in the linear probability model. *International Economic Review*, pages 645–650, 1977.

S. Anderson. Legal Origins and Female HIV. *American Economic Review, forthcoming*, 2018.

S. Anderson and J.-M. Baland. The Economics of ROSCAs and Intrahousehold Resource Allocation. *Quarterly Journal of Economics*, 117(3):963–995, 2002.

N. Ashraf, O. Bandiera, and B. K. Jack. No margin, no mission? A field experiment on incentives for public service delivery. *Journal of Public Economics*, 120:1–17, 2014a.

N. Ashraf, E. Field, and J. Lee. Household bargaining and excess fertility: an experimental study in Zambia. *American Economic Review*, 104(7):2210–2237, 2014b.

N. Ashraf, E. Field, G. Rusconi, A. Voena, and R. Ziparo. Traditional beliefs and learning about maternal risk in Zambia. *American Economic Review*, 107(5):511–515, 2017.

O. P. Attanasio and V. Lechene. Efficient responses to targeted cash transfers. *Journal of Political Economy*, 122(1):178–222, 2014.

S. J. Baird, R. S. Garfein, C. T. McIntosh, and B. Özler. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *The Lancet*, 379(9823):1320–1329, 2012.

O. Bandiera, N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, M. Sulaiman, et al. Women's economic empowerment in action: Evidence from a randomized control trial in Africa. Technical report, International Labour Organization, 2015.

A. Banerjee, D. Karlan, and J. Zinman. Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7 (1):1–21, 2015.

M. Beksinska, J. Smit, C. Joanis, and C. Hart. Practice makes perfect: reduction in female condom failures and user problems with short-term experience in a randomized trial. *Contraception*, 86(2):127–131, 2012.

M. Bjorkman Nyqvist, L. Corno, D. de Walque, and J. Svensson. Using lotteries to incentivize safer sexual behavior: evidence from a randomized controlled trial on HIV prevention. *World Bank Policy Research Working Paper No. 7215*, 2015.

M. Browning and P.-A. Chiappori. Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 66(6):1241–1278, 1998.

M. Bruhn and D. McKenzie. In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1 (4):200–232, 2009.

V. Chernozhukov, I. Fernández-Val, J. Hahn, and W. Newey. Average and quantile effects in nonseparable panel models. *Econometrica*, 81(2):535–580, 2013.

P.-A. Chiappori. Collective labor supply and welfare. *Journal of Political Economy*, 100(3):437–467, 1992.

J. Cohen, K. Lofgren, and M. McConnell. Precommitment, cash transfers, and timely arrival for birth: evidence from a randomized controlled trial in Nairobi Kenya. *American Economic Review*, 107(5):501–05, 2017.

E. N. Coman, M. R. Weeks, I. Yanovitzky, E. Iordache, R. Barbour, M. A. Coman, and T. B. Huedo-Medina. The impact of information about the female condom on female condom use among males and females from a US urban community. *AIDS and Behavior*, 17(6):2194–2201, 2013.

A. Creese, K. Floyd, A. Alban, and L. Guinness. Cost-effectiveness of HIV/AIDS interventions in Africa: a systematic review of the evidence. *The Lancet*, 359(9318): 1635–1642, 2002.

J. Das, J. Hammer, and C. Sánchez-Paramo. The impact of recall periods on reported morbidity and health seeking behavior. *Journal of Development Economics*, 98(1): 76–88, 2012.

D. De Walque, W. H. Dow, R. Nathan, R. Abdul, F. Abilahi, E. Gong, Z. Isdahl, J. Jamison, B. Jullu, S. Krishnan, et al. Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ Open*, 2(1), 2012.

DHS. Mozambique Demographic and Health Survey 2011 – Final report. 2011. URL https://dhsprogram.com/pubs/pdf/FR266/FR266.pdf.

A. Donald, G. Koolwal, J. Annan, K. Falb, and M. Goldstein. Measuring women's agency. *World Bank Policy Research Working Paper No. 8148*, 2017.

D. W. Dowdy, M. D. Sweat, and D. R. Holtgrave. Country-wide distribution of the nitrile female condom (FC2) in Brazil and South Africa: a cost-effectiveness analysis. *AIDS*, 20(16):2091–2098, 2006.

E. Duflo. Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa. *The World Bank Economic Review*, 17(1):1–25, 2003.

E. Duflo and C. Udry. Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices. *NBER Working Paper No. 10498*, 2004.

E. Duflo, P. Dupas, and M. Kremer. Education, HIV, and early fertility: Experimental evidence from Kenya. *American Economic Review*, 105(9):2757–2797, 2015.

P. Dupas. Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*, 3(1):1–34, 2011.

M. Eswaran. The empowerment of women, fertility, and child mortality: towards a theoretical analysis. *Journal of Population Economics*, 15(3):433–454, 2002.

G. Farr, H. Gabelnick, K. Sturgen, and L. Dorflinger. Contraceptive efficacy and acceptability of the female condom. *American Journal of Public Health*, 84(12):1960–1964, 1994.

D. Firth. Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38, 1993.

A. L. Fontanet, J. Saba, V. Chandelying, C. Sakondhavat, P. Bhiraleus, S. Rugpao, C. Chongsomchai, O. Kiriwat, S. Tovanabutra, L. Dally, et al. Protection against sexually transmitted diseases by granting sex workers in Thailand the choice of using the male or female condom: results from a randomized controlled trial. *AIDS*, 12 (14):1851–1859, 1998.

P. Gertler, M. Shah, and S. M. Bertozzi. Risky business: the market for unprotected commercial sex. *Journal of Political Economy*, 113(3):518–550, 2005.

C. Goldin and L. F. Katz. The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, 110(4):730–770, 2002.

J. Greenwood, P. Kircher, C. Santos, and M. Tertilt. An equilibrium model of the African HIV/AIDS epidemic. *NBER Working Paper No. 18953*, 2017.

J. J. Heckman. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. 1981.

G. Heinze and M. Schemper. A solution to the problem of separation in logistic regression. *Statistics in Medicine*, 21(16):2409–2419, 2002.

W. C. Horrace and R. L. Oaxaca. Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3):321–327, 2006.

G. King and L. Zeng. Logistic regression in rare events data. *Political Analysis*, 9(2): 137–163, 2001a.

G. King and L. Zeng. Explaining rare events in international relations. *International Organization*, 55(3):693–715, 2001b.

G. King and L. Zeng. Estimating risk and rate levels, ratios and differences in case-control studies. *Statistics in Medicine*, 21(10):1409–1427, 2002.

W. Koster, M. G. Bruinderink, and W. Janssens. Empowering Women or Pleasing Men? Analyzing Male Views on Female Condom Use In Zimbabwe, Nigeria and Cameroon. *International Perspectives on Sexual and Reproductive Health*, 41(3):126–135, 2015.

T. Lancaster. Orthogonal parameters and panel data. *Review of Economic Studies*, 69 (3):647–666, 2002.

D. S. Lee. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102, 2009.

H. Leitgöb. The problem of modeling rare events in ML-based logistic regression.

M. Macia, P. Maharaj, and A. Gresh. Masculinity and male sexual behaviour in Mozambique. *Culture, Health & Sexuality*, 13(10):1181–1192, 2011.

J. E. Mantell, J. A. Smit, T. M. Exner, Z. Mabude, S. Hoffman, M. Beksinska, E. A. Kelvin, C. Ngoloyi, C.-S. Leu, and Z. A. Stein. Promoting female condom use among female university students in KwaZulu-Natal, South Africa: results of a randomized behavioral trial. *AIDS and Behavior*, 19(7):1129–1140, 2015.

D. McKenzie. Another reason to prefer Ancova: dealing with changes in measurement between baseline and follow-up. *World Bank blog post*, 2015. URL https://blogs.worldbank.org/impactevaluations/another-reason-prefer-ancova-dealing-changes-measurement-between-baseline-and-follow.

G. Miller and A. M. Mobarak. Gender differences in preferences, intra-household externalities, and low demand for improved cookstoves. *NBER Working Paper No. 18964*, 2013.

I. N. d. E. Ministério da Saúde. Inquérito de Indicadores de Imunização Malária e HIV/SIDA em Moçambique: Relatório de Indicadores Básicos de HIV. *DHS Publications*, 2015. URL https://dhsprogram.com/pubs/pdf/PR85/PR85.pdf.

S. Mohapatra and L. Simon. Intra-household bargaining over household technology adoption. *Review of Economics of the Household*, 15(4):1263–1290, 2017.

E. Oster. Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic. *Quarterly Journal of Economics*, 120(2):467–515, 2005.

A. Peters, W. Jansen, and F. van Driel. The female condom: the international denial of a strong potential. *Reproductive Health Matters*, 18(35):119–128, 2010.

A. Philpott, W. Knerr, and D. Maher. Promoting protection and pleasure: amplifying the effectiveness of barriers against sexually transmitted infections and pregnancy. *The Lancet*, 368(9551):2028–2031, 2006.

K. M. Pilz. Assessment of the female condom supply chain in Mozambique. *i+solutions Report*, 2014.

I. Rasul. Household bargaining over fertility: Theory and evidence from Malaysia. *Journal of Development Economics*, 86(2):215–241, 2008.

S. Schaner. Do opposites detract? Intrahousehold preference heterogeneity and inefficient strategic savings. *American Economic Journal: Applied Economics*, 7(2): 135–174, 2015.

A. Schuyler, T. Masvawure, J. Smit, M. Beksinska, Z. Mabude, C. Ngoloyi, and J. Mantell. Building young women's knowledge and skills in female condom use: lessons learned from a South African intervention. *Health Education Research*, 31(2):260–272, 2016.

Y. Stopnitzky. No toilet no bride? Intrahousehold bargaining in male-skewed marriage markets in India. *Journal of Development Economics*, 127:269–282, 2017.

F. Terris-Prestholt and F. Windmeijer. How to sell a condom? The impact of demand creation tools on male and female condom sales in resource limited settings. *Journal of Health Economics*, 48:107–120, 2016.

R. L. Thornton. The demand for, and impact of, learning HIV status. *American Economic Review*, 98(5):1829–1863, 2008.

J. Trussell. Contraceptive failure in the United States. *Contraception*, 83:397–404, 2011.

UNAIDS. UNAIDS Prevention Gap Report. 2016a. URL http://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf.

UNAIDS. HIV prevention among adolescent girls and young women. 2016b. URL http://www.unaids.org/sites/default/files/media_asset/UNAIDS_HIV_prevention_among_adolescent_girls_and_young_women.pdf.

E. Van Tassel. Household bargaining and microfinance. *Journal of Development Economics*, 74(2):449–468, 2004.

G. Vijayakumar, Z. Mabude, J. Smit, M. Beksinska, and M. Lurie. A review of female-condom effectiveness: patterns of use and impact on protected sex acts and STI incidence. *International Journal of STD & AIDS*, 17(10):652–659, 2006.

R. P. Walensky, E. L. Ross, N. Kumarasamy, R. Wood, F. Noubary, A. D. Paltiel, Y. M. Nakamura, S. V. Godbole, R. Panchia, I. Sanne, et al. Cost-effectiveness of HIV treatment as prevention in serodiscordant couples. *New England Journal of Medicine*, 369(18):1715–1725, 2013.

R. K. Wanyenze, L. Atuyambe, V. Kibirige, S. Mbabazi, N. M. Tumwesigye, K. Djurhuus, and A. Namale. The new female condom (FC2) in Uganda: perceptions and experiences of users and their sexual partners. *African Journal of AIDS Research*, 10(3):219–224, 2011.

A. Young. Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Working Paper*, 2016.

A Appendix

Table A.1: Treatment effects – primary outcome variables – inverse probability weighting

	(1) **Ever FC**	(2) **Ever MC**	(3) **Ever other**	(4) **Last 30 days FC**	(5) **Last 30 days MC**	(6) **Current FC**	(7) **Current MC**	(8) **Current 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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to 8 to whether she is currently using it. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors 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 STI and HIV status which have missing values which are likely non-random.

Table A.2: Lee bounds – primary outcome variables

	(1) **Ever FC**	(2) **Ever MC**	(3) **Ever other**	(4) **Last 30 days FC**	(5) **Last 30 days MC**	(6) **Current FC**	(7) **Current MC**	(8) **Current 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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to whether she is currently using it. Bounds do not include facilitator fixed effects, as attrition is not monotonic on treatment status conditional on facilitator fixed effects.

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 $u_i(P, H) \geq u_i^0 \;\; \forall i = m, f$. Define

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

as the set of all points (P, H) that satisfies both partner's participation constraints. [24]

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

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

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

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

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

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

[24]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

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

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

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

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

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

which yields

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

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

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

and thus that

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

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

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

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

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

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 A.1.

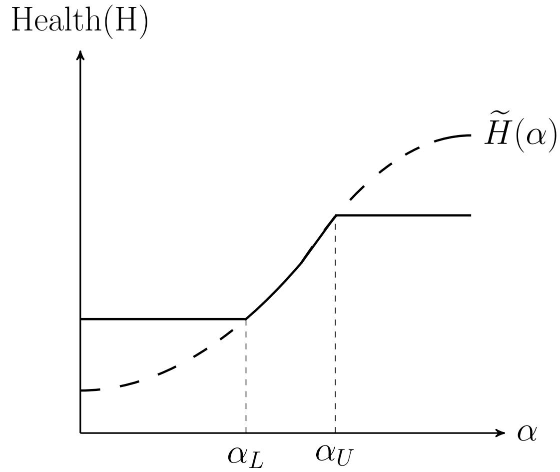
When only the binary set $\{US, MC\}$ is available, it follows directly from the weakly increasing nature of $H^*\left(\alpha\right)$ that there will be cut-off values of α such that

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

The introduction of female condoms expands the available technologies to the ternary set $\{US, FC, MC\}$.[25] Given that $H_{MC} > H_{FC} > H_{US}$, it follows directly that there

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

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



will threshold values of α such that

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

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

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

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

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

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

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

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

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

and

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

implying a solution $\widetilde{q}_j(\alpha)$ for $j = m, f$. Note that $\phi'(q) \leq 1)$ 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

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

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

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

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

with the following Lagrangean

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

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 A.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

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

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

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

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

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

with Lagrangean

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

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 A.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 A.3: 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	-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
HH 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	-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
Names FC as contraceptive	-0.04	0.53	-0.06	0.43	0.00	0.97	150	146	296
Baseline use									
Ever used FC	0.05	0.60	0.12	0.32	0.07	0.80	152	146	298
Ever used MC	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 MC last 30 days	-0.04	0.53	-0.07	0.42	0.01	0.92	152	146	298
Current use MC	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 prior to attrition. Lower sample sizes reflect observations that are missing or not applicable. "Treatment" contains all individuals assigned to the treatment group, whether or not they attended the 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. "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 A.4: Baseline balance excluding attritters

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age	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
HH 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	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
STD last 3 months (self-report)	0.14	0.15	0.12	0.58	205	92	113
Names FC as contraceptive	0.43	0.46	0.40	0.91	230	107	123
Contraceptive use							
Ever used FC	0.08	0.07	0.08	-0.15	232	107	125
Ever used MC	0.74	0.77	0.71	0.94	232	107	125
Ever used other	0.72	0.71	0.74	-0.43	232	107	125
Never used any	0.07	0.08	0.06	0.58	232	107	125
Used FC last 30 days	0.03	0.02	0.04	-0.97	232	107	125
Used MC last 30 days	0.33	0.30	0.36	-0.98	232	107	125
Current use FC	0.03	0.03	0.03	-0.18	232	107	125
Current use MC	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
Current use none	0.37	0.38	0.35	0.49	232	107	125

Notes: N=232 in the balanced sample excluding attritters. Lower sample sizes reflect observations that are missing or not applicable. “Treatment Mean” contains all individuals assigned to the treatment group, whether or not they attended the 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.

Table A.5: 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

Factor loadings from a PCA of all asset questions simultaneously. Only loadings greater than or equal to 0.25 are displayed. All variables are coded such that a higher value proxies greater bargaining power for the respondent. As shown, the asset questions load three separate components, all of which are used in the heterogeneity analysis.

Table A.6: Bargaining power – principal component analysis

	(PC1)	(PC2)	Unexplained
Decision-making (1= she is involved)			
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 influence			0.8677
In general, more power			0.9097
Power dynamics (1=completely disagree)			
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 PCA of 'decision-making' and 'power dynamics' bargaining power variables simultaneously. Only loadings greater than or equal to 0.25 are displayed. All variables are coded such that a higher value proxies greater bargaining power for the respondent.

Table A.7: 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. "Assets 1", "Assets 2" and "Assets 3" represent the three principal components loaded by the assets module as identified in Table A.5. "Decision-making" and "Power dynamics" represent the principal component loaded by each of these survey modules as identified in Table A.6. All variables are 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. "Treatment Mean" contains all individuals assigned to the treatment group, whether or not they attended the 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. All values taken from the baseline survey (N=298). The assets module 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). The decision-making module was enumerated to all respondents (N=298), except the questions "who has more influence" and "who has more power" which were asked only of women in a stable relationship at baseline (N=250). Power dynamics questions were 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.

Table A.8: Baseline balance on covariates – diary subsample

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Demographics							
Age	31.32	31.93	30.67	0.55	56	29	27
Education	5.95	5.45	6.48	-1.32	56	29	27
Literate	0.84	0.86	0.81	0.42	55	28	27
HH 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	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
STD last 3 months (self-report)	0.12	0.12	0.12	0.00	48	24	24
Names FC as contraceptive	0.27	0.31	0.23	0.66	55	29	26

Notes: N=56 in the balanced panel. Lower sample sizes reflect observations that are missing or not applicable. "Treatment Mean" contains all individuals assigned to the treatment group, whether or not they attended the sessions. Column 4 presents the t-test statistic for the null hypothesis that the mean in the treatment group is equal to the mean in the control group.

Table A.9: Baseline balance on use – diary subsample

	Mean	Control Mean	Treatment Mean	t-test	Total N	Control N	Treatment N
Ever use survey							
Ever used FC	0.05	0.07	0.04	0.53	56	29	27
Ever used MC	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 FC last 30 days	0.04	0.00	0.07	-1.44	56	29	27
Used MC last 30 days	0.36	0.28	0.44	-1.31	56	29	27
Current use survey							
Current use FC	0.02	0.00	0.04	-1.00	56	29	27
Current use MC	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
Current use none	0.27	0.34	0.19	1.36	56	29	27
Ever used diaries							
Used FC in baseline weeks	0.04	0.03	0.04	-0.05	56	29	27
Used MC in baseline weeks	0.45	0.41	0.48	-0.50	56	29	27
Use last 30 days							
Used FC in baseline last 30 days	0.04	0.03	0.04	-0.05	56	29	27
Used MC in baseline last 30 days	0.54	0.52	0.56	-0.28	56	29	27
Use last 14 days							
Used FC in baseline last 14 days	0.04	0.03	0.04	-0.05	56	29	27
Used MC in baseline last 14 days	0.39	0.34	0.44	-0.75	56	29	27
Impact on discussions and sex acts							
% 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 in the balanced panel. Lower sample sizes reflect observations that are missing or not applicable. "Treatment Mean" contains all individuals assigned to the treatment group, whether or not they attended the 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. Percentage use represents the percentage of sex acts per individual which were protected by female condoms, male condoms and other contraceptives respectively.

Table A.10: Diary sample representativeness of full sample – covariates

	Mean	Survey Mean	Diary subsample Mean	t-test	Survey N	Diary subsample N
Demographics						
Age	30.95	30.74	31.92	-0.90	231	50
Years of education	6.21	6.30	5.76	1.17	230	50
Literate	0.85	0.86	0.82	0.73	229	49
HH head	0.25	0.23	0.34	-1.45	231	50
Income						
Has job	0.38	0.37	0.42	-0.61	228	50
Personal income last 30 days	871.71	816.67	1126.00	-1.27	231	50
Relationships						
In a stable relationship (incl. married)	0.85	0.85	0.84	0.22	231	50
Married (officially or unofficially)	0.60	0.62	0.50	1.50	230	50
Years relation	9.70	9.20	12.19	-1.97*	183	37
# Partners last 12 months	0.93	0.93	0.92	0.20	231	50
Sexual behaviour						
Pregnant	0.05	0.06	0.00	3.85***	230	50
HIV positive (self-report)	0.31	0.30	0.35	-0.56	201	43
STI last 3 months (self-report)	0.13	0.14	0.12	0.38	204	43
Names FC as contraceptive	0.41	0.43	0.31	1.64	229	49
Bargaining power (principle components)						
Assets 1	0.79	0.80	0.78	0.09	204	43
Assets 2	0.48	0.48	0.47	0.05	204	43
Assets 3	1.21	1.23	1.14	0.70	204	43
Decision-making	1.74	1.69	1.98	-1.58	160	34
Power dynamics	2.92	2.93	2.87	0.36	160	34

Notes: N=232 im the baseline sample of which N=57 are in the subsample who respond to the diaries. Lower sample sizes reflect observations that are missing or not applicable. "Survey Mean" contains all individuals in the balanced panel, whether or not they participated in the diaries. "Diary subsample Mean" contains just those individuals who 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 full sample.

Table A.11: Diary sample representativeness of full sample – baseline use

	Mean	**Survey Mean**	**Diaries Mean**	**t-test**	**Survey N**	**Diaries N**
Ever use survey						
Ever used FC	0.08	0.09	0.05	0.94	297	55
Ever used MC	0.75	0.75	0.78	-0.56	297	55
Ever used other	0.74	0.72	0.85	-2.45**	297	55
Use last 30 days survey						
Used FC last 30 days	0.03	0.03	0.04	-0.35	297	55
Used MC last 30 days	0.32	0.32	0.36	-0.67	297	55
Current use survey						
Current use FC	0.02	0.02	0.02	0.27	297	55
Current use MC	0.39	0.39	0.40	-0.08	297	55
Current use other	0.41	0.39	0.49	-1.36	297	55

Notes: N=57 are in the subsample of survey respondents who respond to the diaries. N=298 are in the full baseline study sample. Lower sample sizes reflect observations that are missing. All values taken from the baseline survey. "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. 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 full sample.

B.3 Additional Analyses

Table A.12: Treatment effects – heterogeneity by relationship status

	(1) **Ever FC**	(2) **Ever MC**	(3) **Last 30 days FC**	(4) **Last 30 days MC**	(5) **Current FC**	(6) **Current MC**
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. 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. Dependent variables are binary indicators for the current use of female and male condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table A.13: Current use of female condoms at endline – correlation with bargaining power among women using male condoms at baseline

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. The sample is restricted to women who report currently using male condoms at baseline, and who answer the assets module (columns 1-3) or all two survey modules on bargaining power (columns 4-5). "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table A.5. "Decision-making" and"Power dynamics" are the first two principal components, as identified in Table A.6. Dependent variables are binary indicators for whether the respondent is currently using a female condom at endline. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. 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. 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 or were not using condoms at baseline.

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 A.21 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 29.[26] Table A.22 show the results. 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.[27] 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.

[26] Again, similar results are obtained with a logit specification (see A.37).

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

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

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 A.14: Impacts on current use of female condoms – LASSO selects heterogeneity by bargaining power

	(1)	(2)	(3)	(4)	(5)
	β / (s.e.)	β / (s.e.)	β / (s.e.)	β / (s.e.)	β / (s.e.)
Treatment	0.446***	0.384**	0.427***	0.550***	0.506
	(0.161)	(0.156)	(0.160)	(0.181)	(0.199)
Assets 1	-0.009				
	(0.007)				
Treat*Assets1	-0.038**				
	(0.016)				
Assets 2					
Treat*Assets2					
Assets 3					
Treat*Assets3			-0.056***		
			(0.020)		
Decision-making					
Treat*Decision				-0.104***	
				(0.036)	
Power dynamics					0.019
					(0.018)
Treat*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)	
Treat*Age	-0.008**	-0.007**	-0.007**	-0.006*	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Treat*Years of education	-0.011	-0.012	-0.008	-0.011	-0.007
	(0.011)		(0.011)	(0.009)	(0.010)
Treat*Literacy	0.052	0.058	0.055	0.032	0.019
	(0.111)	(0.114)	(0.111)	(0.112)	(0.128)
Treat*high HIV risk perception	- 0.142*	-0.135*	-0.141*	-0.134*	-0.107
	(0.080)	(0.100)	(0.080)	(0.080)	(0.079)
Treat*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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Excluding attritters, N=198 women are in a stable relationship at baseline. N=206 answer the assets module, including some women who live with their partner but do not report being in a stable relationship. N=160 have non-missing values for all of the decision-making and power dynamics questions (which is required for the principal components to be calculated) and all of the controls. "Treatment" is a dummy for being assigned to the treatment group. 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 identified in Table A.5. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table A.6. Dependent variables are binary indicators for current use of female condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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.

Table A.15: Impacts on current use of male condoms – heterogeneity by bargaining power

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
Treatment	0.017	0.056	-0.009	0.040	0.138
	(0.076)	(0.065)	(0.106)	(0.142)	(0.181)
Assets 1	-0.060*				
	(0.031)				
Treat×Assets1	0.092*				
	(0.055)				
Assets 2		-0.045			
		(0.031)			
Treat×Assets2		0.076*			
		(0.040)			
Assets 3			-0.045*		
			(0.027)		
Treat×Assets3			0.082		
			(0.077)		
Decision-making				0.010	
				(0.049)	
Treat×Decision				0.008	
				(0.064)	
Power dynamics					0.043
					(0.043)
Treat×Power dynamics					-0.029
					(0.060)
Controls	✓	✓	✓	✓	✓
Observations	198	198	198	180	180
Control mean endline	0.353	0.353	0.353	0.353	0.353

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Excluding attritters, N=198 women are in a stable relationship at baseline. N=206 answer the assets module, including some women who live with their partner but do not report being in a stable relationship. N=160 have non-missing values for all of the decision-making and power dynamics questions (which is required for the principal components to be calculated) and all of the controls. "Treatment" is a dummy for being assigned to the treatment group. 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 identified in Table A.5. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table A.6. Dependent variables are binary indicators for current use of male condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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.

Table A.16: Treatment effects – heterogeneity by baseline bargaining power, controlling for baseline use of other contraceptives

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
Treatment	0.123**	0.109**	0.153***	0.207**	0.173
	(0.050)	(0.047)	(0.055)	(0.099)	(0.108)
Current use other (baseline)	-0.001	-0.002	-0.003	-0.026	-0.031
	(0.037)	(0.038)	(0.036)	(0.027)	(0.029)
Treat×Current use other (baseline)	0.014	0.005	0.016	0.036	0.058
	(0.077)	(0.078)	(0.077)	(0.075)	(0.078)
Treat×Assets1	-0.045**				
	(0.018)				
Assets 1					
Treat×Assets2		-0.031**			
		(0.016)			
Assets 2					
Treat×Assets3			-0.055***		
			(0.019)		
Assets 3					
Treat×Decision				-0.078**	
				(0.039)	
Decision-making					
Treat×Power dynamics					-0.038
					(0.040)
Power dynamics					
Controls	✓	✓	✓	✓	✓
Observations	198	198	198	180	180
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. N=198 women are in a stable relationship at baseline. N=160 have non-missing values for all of the decision-making and power dynamics questions (required for the principal components to be calculated) and all of the controls. N=206 answer the assets module, including some women who live with their partner but do not report being in a stable relationship. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table A.6."Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table A.5. Dependent variables are binary indicators for the current use of female condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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. "Current use other (baseline)" is a dummy equal to one if the respondent reports using any non-condom forms of modern contraception at baseline, which mainly comprises either the pill or injectables. 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.

Table A.17: Treatment effects – heterogeneity by baseline bargaining power, controlling for HIV status

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
Treatment	0.089*	0.076*	0.120**	0.224**	0.205
	(0.051)	(0.043)	(0.058)	(0.089)	(0.136)
HIV positive (self-report)	0.017	0.020	0.020	0.046	0.046
	(0.042)	(0.041)	(0.042)	(0.045)	(0.046)
Treatment×HIV positive	0.084	0.086	0.083	0.066	0.034
	(0.098)	(0.098)	(0.099)	(0.101)	(0.104)
Treat×Assets1	-0.034*				
	(0.019)				
Assets 1					
Treat×Assets2		-0.024			
		(0.015)			
Assets 2					
Treat×Assets3			-0.048**		
			(0.019)		
Assets 3					
Treat×Decision				-0.087**	
				(0.042)	
Decision-making					
Treat×Power dynamics					-0.042
					(0.046)
Power dynamics					
Controls	✓	✓	✓	✓	✓
Observations	170	170	170	156	156
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. N=198 women are in a stable relationship at baseline. N=160 have non-missing values for all of the decision-making and power dynamics questions (required for the principal components to be calculated) and all of the controls. N=206 answer the assets module, including some women who live with their partner but do not report being in a stable relationship. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table A.6. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table A.5. Dependent variables are binary indicators for the current use of female condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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. "HIV positive (self-report)" is a dummy equal to one if the respondent reports being HIV-positive at basline. 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.

Table A.18: Treatment effects – heterogeneity by baseline bargaining power, controlling for risk beliefs

	(1) β / (s.e.)	(2) β / (s.e.)	(3) β / (s.e.)	(4) β / (s.e.)	(5) β / (s.e.)
Treatment	0.249***	0.223***	0.279***	0.341***	0.268**
	(0.086)	(0.081)	(0.090)	(0.120)	(0.125)
Believes high risk to self	-0.029	-0.030	-0.035	-0.002	-0.007
	(0.042)	(0.043)	(0.042)	(0.035)	(0.038)
Treatment×Believes high risk to self	-0.151*	-0.145*	-0.147*	-0.137*	-0.101
	(0.081)	(0.081)	(0.081)	(0.075)	(0.079)
Treat×Assets1	-0.050**				
	(0.020)				
Assets 1	0.000				
	(0.011)				
Treat×Assets2		-0.033**			
		(0.016)			
Assets 2		0.008			
		(0.011)			
Treat×Assets3			-0.061***		
			(0.021)		
Assets 3			0.001		
			(0.005)		
Treat×Decision				-0.090**	
				(0.038)	
Decision-making				-0.015	
				(0.021)	
Treat×Power dynamics					-0.037
					(0.041)
Power dynamics					0.014
					(0.020)
Controls	✓	✓	✓	✓	✓
Observations	197	197	197	179	179
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. N=198 women are in a stable relationship at baseline. N=160 have non-missing values for all of the decision-making and power dynamics questions (required for the principal components to be calculated) and all of the controls. N=206 answer the assets module, including some women who live with their partner but do not report being in a stable relationship. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. "Decision-making" and "Power dynamics" are the first two principal component from each of these modules, as identified in Table A.6. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table A.5. Dependent variables are binary indicators for the current use of female condoms. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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. "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". 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.

Table A.19: 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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. 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). N=232 for the endline survey excluding attritters, except the violence module where N=162 since these questions were only enumerated to women in a stable relationship. Missing observations reflect not applicable, does not want to answer, and cases where the facilitator fixed effect perfectly predicts the outcome variable. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator fixed effects (N=17), since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table A.20: 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: "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are bargaining power indicators measured at endline. All regressions are linear probability model ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All regressions include facilitator fixed effects (N=17), since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation. N=232 for the endline survey excluding attritters.

Table A.21: Treatment effects – survey variables, diary subsample

	(1) **Ever FC**	(2) **Ever MC**	(3) **Ever other**	(4) **Last 30 days FC**	(5) **Last 30 days MC**	(6) **Current FC**	(7) **Current MC**	(8) **Current 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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for the use of contraceptives at the level of the respondent. "Treatment" is an indicator for observations in the treatment group. 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 models with the respondent as the unit of observation. Facilitator fixed effects are dropped because facilitator perfectly predicts outcomes for many observations in this subsample. Standard errors are clustered at the individual level, since this was the level of randomisation.

Table A.22: Impacts on condom use – fixed effects panel estimator, diary subsample

	(1) **FC** **full endline period** β / (s.e.)	(2) **MC** **full endline period** β / (s.e.)	(3) **FC** **last 30 days** β / (s.e.)	(4) **MC** **last 30 days** β / (s.e.)	(5) **FC** **last 14 days** β / (s.e.)	(6) **MC** **last 14 days** β / (s.e.)
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)
Individual f.e.'s	✓	✓	✓	✓	✓	✓
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

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for the proportion of sex acts that protected by male and female condom use per week per respondent. Weeks with 0 sex acts are counted as missing. "Treat×endline" is an indicator for observations in the treatment group during the endline period. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat×endline" is the intent-to-treat effect. All regressions are respondent level OLS fixed effects models with the respondent-week as the unit of observation. All specifications include individual fixed effects (N=57) and facilitator×endline fixed effects (N=17) for inference since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

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:

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

where Y_{jift} is the outcome variable of interest relating to sex act j, reported by individual i, who was assigned to facilitator f.[28] *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 A.23 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.

[28]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 A.23: Impacts on likelihood that a sex act is protected – diary subsample

	(1) **FC-protected full endline period** β / (s.e.)	(2) **MC-protected full endline period** β / (s.e.)	(3) **FC-protected last 30 days** β / (s.e.)	(4) **MC-protected last 30 days** β / (s.e.)	(5) **FC-protected last 14 days** β / (s.e.)	(6) **MC-protected last 14 days** β / (s.e.)
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 f.e.'s	✓	✓	✓	✓	✓	✓
Observations	349	349	204	204	143	143
Control mean	0.010	0.330	0.009	0.374	0.013	0.387

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. All regressions are linear probability models with the sex act as the unit of observation. Given that by definition there are no repeated baseline and endline observations at the level of the sex act, baseline values do not exist and so are not included as regressors. Dependent variables are binary indicators that take the value of 1 if the sex act was protected by female or male condoms. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat×endline" is the intent-to-treat effect. All specifications include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

Table A.24: Likelihood of sex acts: Interaction of treatment and baseline MC use

	(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 × MCbaseline	-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

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for whether a respondent had a sex act in a particular week. MCbaseline is a binary indicator for whether a respondent currently uses MC at baseline. "Treat × endline" is an indicator for observations in the treatment group during the endline period. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat × endline" is the intent-to-treat effect. All regressions are respondent-level LPM fixed effects models with the respondent-week as the unit of observation. All specifications include individual fixed effects (N=56), and facilitator × endline fixed effects (N=17) for inference since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

Table A.25: Impacts on discussions per respondent per week – diary subsample

	(1) **Discussion full endline period** β / (s.e.)	(2) **Female-initiated full endline period** β / (s.e.)	(3) **Discussion last 30 days** β / (s.e.)	(4) **Female-initiated last 30 days** β / (s.e.)	(5) **Discussion last 14 days** β / (s.e.)	(6) **Female-initiated last 14 days** β / (s.e.)
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
Clusters	50	50	49	49	47	47
Control mean	0.227	0.192	0.275	0.228	0.311	0.265

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for discussions and female-initiated discussions about condom use at the sex-act level, observed per respondent per week. "Treat×endline" is an indicator for observations in the treatment group during the endline period. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat×endline" is the intent-to-treat effect. All regressions are respondent level OLS fixed effects models with the sex-act as the unit of observation. All specifications include individual fixed effects (N=56), and facilitator×endline fixed effects (N=17) for inference since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

Table A.26: Impacts on condom use – ANCOVA specification, diary subsample

	(1) **FC** **full endline period** β / (s.e.)	(2) **MC** **full endline period** β / (s.e.)	(3) **FC** **last 30 days** β / (s.e.)	(4) **MC** **last 30 days** β / (s.e.)	(5) **FC** **last 14 days** β / (s.e.)	(6) **MC** **last 14 days** β / (s.e.)
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	57	57	57	57	57	57
Control mean	0.069	0.552	0.034	0.448	0.034	0.448

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms and male condoms. Columns 1 and 2 refer to whether the respondent reports using that type of condom at least once during the full endline period in the diaries, columns 3 and 4 to whether she reports using it at least once in the last 30 days of the diaries, and columns 5 and 6 to whether she reports using it at least once in the last two weeks of the diaries. 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.

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.[29] 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.[30] 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, and the number of mothers requiring

[29]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.

[30]The SPECTRUM suite is developed by Avenir Health, see http://www.avenirhealth.org/software-spectrum.php.

Prevention of Mother-To-Child Transmission for the period from 2015-2030. 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; see Section B.5 for details.

On the cost side, we 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. 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.

Comparing the programme costs to the DALYs averted allows 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; Oster, 2005)). Comparing the programme costs to the cost savings 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 one the ICER for the full intervention is -50 USD[31], i.e. a saving of 50 USD per DALY averted, meaning that scaling up the full intervention is therefore *very cost-*

[31] 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 the GDP per capita (Walensky et al., 2013) The GDP per capita of Mozambique was 511 USD in 2014.

effective. 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 two 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 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 A.2: Simulation of annual number of HIV infections

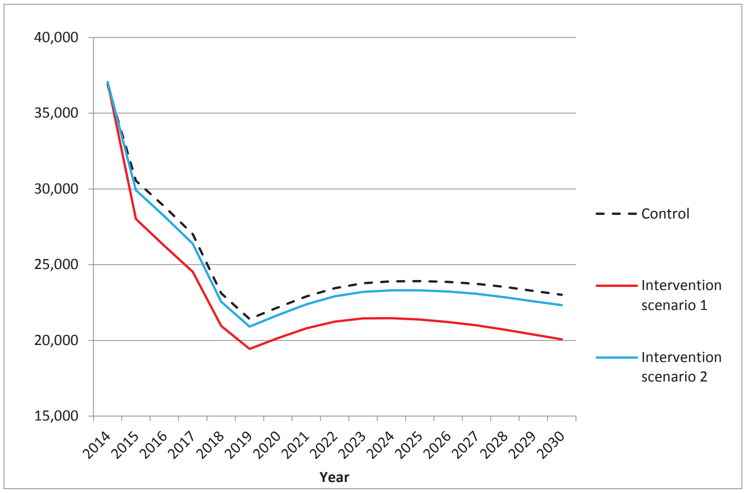


Table A.27: Simulation of impact on HIV infections and DALYs averted by 2030

	# **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) in 2014 was 3,048,905. The statistics are calculated by comparing control projections up to 2030 without any changes to the demographic and behavioural data (control) with interventions projections where behavioural data (condom use) and epidemiological data (condom efficacy) are changed from 2015 onward.

Table A.28: Programme and condom unit and distribution costs – Scenario 1

Year	Female population to be treated	Programme costs intervention (USD)	# of additional condoms	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. Female population to be treated in 2015 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 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 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 is calculated by multiplying the discounted unit and distribution cost of female condoms (0.45 USD in 2015, discounted at 3% per year) with the # of additional condoms required.

Table A.29: Programme and condom unit and distribution costs – Scenario 2

Year	Female population to be treated	Programme costs intervention (USD)	# of additional condoms	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. Female population to be treated in 2015 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 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 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 is calculated by multiplying the discounted unit and distribution cost of female condoms (0.45 USD in 2015, discounted at 3% per year) with the # of additional condoms required.

Table A.30: 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 ART and PMTCT is calculated by making projections in the control and intervention scenario of the number of individuals needing ART and PMTCT. The number of people receiving ART and PMTCT are then calculated by multiplying the number of individuals needing ART 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 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 A.31: 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 ART and PMTCT is calculated by making projections in the control and intervention scenario of the number of individuals needing ART and PMTCT. The number of people receiving ART and PMTCT are then calculated by multiplying the number of individuals needing ART 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 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 A.32: 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
Lab 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)	
ChildrenARVDrugs	124.48
ChildrenLabCostsARTTr	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 visit	1.00
OI treatment: in-patient days	0.00
OI treatment: out-patient days	1.00

Notes: Data for South Mozambique 2015. Based on the SPECTRUM Resource Needs Model

Table A.33: Unit cost PMTCT for mothers Mozambique 2015

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

Notes: Data for South Mozambique 2015. Based on the SPECTRUM Resource Needs Model

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:

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

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 4 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.[32] 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).[33] 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

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

[33] 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).[34] 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

[34]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 4, A.34, and A.35 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 6 and A.37 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 prefect 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 randomisation. 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 A.3. These predicted marginal effects are based on the estimates from the rare events logit (see Table A.38). The dashed red line is the average "intent-to-treat" effect. The blue line 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 A.34: Main treatment outcomes – Logit specification

	(1) **Ever FC** Mfx / (s.e.)	(2) **Ever MC** Mfx / (s.e.)	(3) **Ever other** Mfx / (s.e.)	(4) **Last 30 days FC** Mfx / (s.e.)	(5) **Last 30 days MC** Mfx / (s.e.)	(6) **Current FC** Mfx / (s.e.)	(7) **Current MC** Mfx / (s.e.)	(8) **Current other** Mfx / (s.e.)
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to 8 to whether she is currently using it. All regressions are logit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table A.35: Main treatment outcomes – Firthlogit specification

	(1) **Ever FC** Mfx / (s.e.)	(2) **Ever MC** Mfx / (s.e.)	(3) **Ever other** Mfx / (s.e.)	(4) **Last 30 days FC** Mfx / (s.e.)	(5) **Last 30 days MC** Mfx / (s.e.)	(6) **Current FC** Mfx / (s.e.)	(7) **Current MC** Mfx / (s.e.)	(8) **Current other** Mfx / (s.e.)
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to 8 to whether she is currently using it. All regressions are firthlogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table A.36: Main treatment outcomes – relogit specification

	(1) **Ever FC** Mfx / (s.e.)	(2) **Ever MC** Mfx / (s.e.)	(3) **Ever other** Mfx / (s.e.)	(4) **Last 30 days FC** Mfx / (s.e.)	(5) **Last 30 days MC** Mfx / (s.e.)	(6) **Current FC** Mfx / (s.e.)	(7) **Current MC** Mfx / (s.e.)	(8) **Current other** Mfx / (s.e.)
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: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. "Treatment" is a dummy for being assigned to the treatment group. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treatment" is the intent-to-treat effect. Dependent variables are binary indicators for the use of female condoms, male condoms and other modern contraceptive methods. 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 to 8 to whether she is currently using it. All regressions are relogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. All include facilitator fixed effects (N=17) since randomisation was stratified on facilitator. Standard errors are robust to individual-level heteroskedasticity, since this was the level of randomisation.

Table A.37: Impacts on likelihood of sex acts per respondent week – Logit

	(1) **# Sex acts per week full endline period** β / (s.e.)	(2) **# Sex acts per week last 30 days** β / (s.e.)	(3) **# Sex acts per week last 14 days** β / (s.e.)
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

Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Dependent variables are binary indicators for whether a respondent had a sex act in a particular week. "Treat×endline" is an indicator for observations in the treatment group during the endline period. Not all respondents assigned to treatment attended the sessions, thus the coefficient on "Treat×endline" is the intent-to-treat effect. All regressions are respondent-level logit fixed effects models with the respondent-week as the unit of observation. All specifications include individual fixed effects (N=56), and facilitator×endline fixed effects (N=17) for inference since randomisation was stratified on facilitator. Standard errors are clustered at the level of the individual, since this was the level of randomisation.

Table A.38: Impacts on current use of female condoms – heterogeneity by bargaining power – Relogit

	(1) Mfx / (s.e.)	(2) Mfx / (s.e.)	(3) Mfx / (s.e.)	(4) Mfx / (s.e.)	(5) Mfx / (s.e.)
Treatment	28471757.599***	5.063***	21.857***	1.540	6.300**
	(0.763)	(0.850)	(0.983)	(1.218)	(2.859)
Assets 1	70020910.740***				
	(0.791)				
Treat×Assets1	-7.002e+07				
	(0.000)				
Assets 2		23.603***			
		(1.523)			
Treat×Assets2		-20.773***			
		(1.753)			
Assets 3			21.840***		
			(0.404)		
Treat×Assets3			-22.168***		
			(0.593)		
Decision-making				0.452	
				(0.519)	
Treat×Decision				-1.500**	
				(0.669)	
Power dynamics					1.308*
					(0.675)
Treat×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

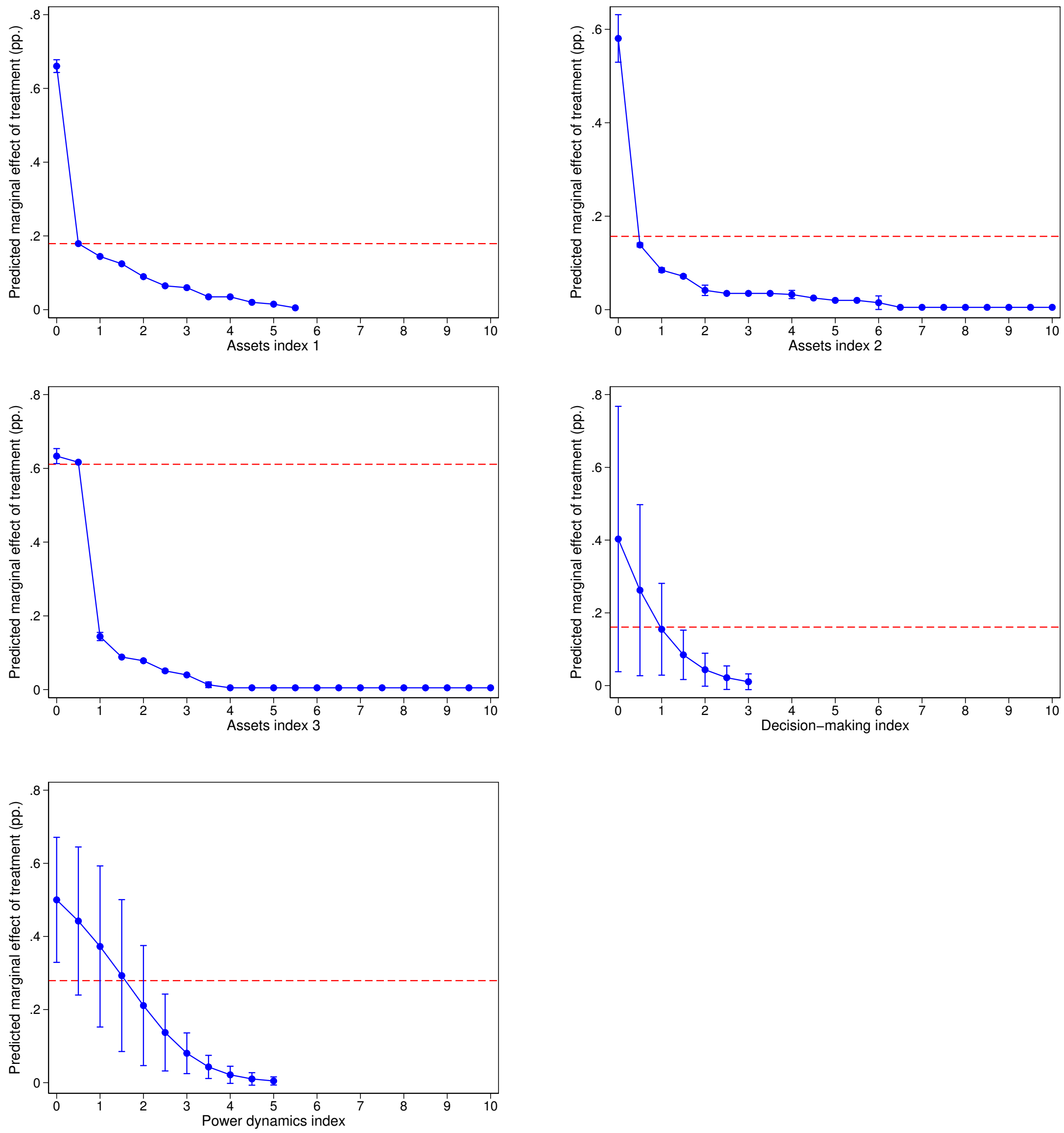
Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Excluding attritters, N=206 women are in a stable relationship at baseline. N=198 are women who are in a stable relationship and have no missing values on the control variables. N=180 have non-missing values for all of the decision-making, power dynamics, and control variables. "Treatment" is a dummy for being assigned to the treatment group. 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 identified in Table A.5. "Decision-making" and "Power dynamics" are the first two principal component from all the survey questions referring to these two modules, as identified in Table A.6. Dependent variables are binary indicators for current use of female condoms. All regressions are relogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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.

Table A.39: Impacts on current use of female condoms – heterogeneity by bargaining power – Firthlogit

	(1) Mfx / (s.e.)	(2) Mfx / (s.e.)	(3) Mfx / (s.e.)	(4) Mfx / (s.e.)	(5) Mfx / (s.e.)
Treatment	0.132**	0.114**	0.172**	0.096	0.452**
	(0.067)	(0.046)	(0.083)	(0.076)	(0.116)
Assets 1	0.014				
	(0.071)				
Treat×Assets1	-0.098				
	(0.146)				
Assets 2		0.045			
		(0.048)			
Treat×Assets2		-0.056			
		(0.065)			
Assets 3			0.030		
			(0.027)		
Treat×Assets3			-0.113		
			(0.128)		
Decision-making				-0.087	
				(0.064)	
Treat×Decision				-0.009	
				(0.104)	
Power dynamics					0.162**
					(0.071)
Treat×Power dynamics					-0.177***
					(0.059)
Controls	✓	✓	✓	✓	✓
Observations	198	198	198	182	182
Control mean endline	0.020	0.020	0.020	0.020	0.020

Notes: Significance levels $p < 0.10^{*}$, $p < 0.05^{**}$, $p < 0.01^{***}$. Excluding attritters, N=206 women are in a stable relationship at baseline. N=198 are women who are in a stable relationship and have no missing values on the control variables. N=180 have non-missing values for all of the decision-making, power dynamics, and control variables. "Treatment" is a dummy for being assigned to the treatment group. 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 identified in Table A.5. "Decision-making" and "Power dynamics" are the first two principal component from all the survey questions referring to these two modules, as identified in Table A.6. Dependent variables are binary indicators for current use of female condoms. All regressions are Firthlogit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects due to loss of sample size where baseline use perfectly predicts endline use conditional on a given facilitator. Standard errors are not clustered since this is not an option with Firthlogit. 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.

Figure A.3: Treatment effect on current use of FC – Heterogeneity by bargaining power indices – Relogit



Notes: Predicted marginal effects of treatment on the outcome variable "respondent is currently using female condoms at endline", estimated with rare events logit (see Table A.38), with 95% confidence intervals. Effects are estimated from the interaction of treatment with the bargaining power indices. Indices are constructed via PCA of the relevant survey modules, and are re-centered such that zero is the minimum and one unit represents one standard deviation. The treatment indicator is equal to one if the respondent was assigned to the treatment group, thus the estimates are "intent-to-treat" effects. The dashed red line is the average intent-to-treat effect. "Assets 1", "Assets 2" and "Assets 3" are the first three principal components from the assets module, as identified in Table A.5. "Decision-making" and "Power dynamics" are the first two principal component from all the survey questions referring to these two modules, as identified in Table A.6. All regressions are rare events logit ANCOVA specifications, including the baseline value of the dependent variable as a regressor. Regressions do not include facilitator fixed effects 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.