# Family Planning and Sexual Well-Being: Experimental Evidence from Urban Malawi

Mahesh Karra[a] Sarah Vincent[b] David Canning[c]

August 14, 2026

**Abstract**

Family planning (FP) programs have been historically centered on averting negative reproductive outcomes, such as unintended pregnancies and unwanted births, even as these same programs call for the equal promotion of positive reproductive outcomes, in particular sexual health and well-being. In this study, we examine the causal effect of FP on sexual well-being using evidence from a randomized controlled trial in urban Malawi. Married women aged 18 to 35 who were either pregnant or had given birth within the previous six months were recruited and randomly assigned to an intervention arm or a pure control arm. Women who were assigned to the intervention arm received a multicomponent FP intervention over a two year period. On the extensive margin, intervention women initiate sex sooner but not necessarily more often: the FP intervention raises the likelihood of recent sexual activity but not coital frequency among already sexually active women. Among women who had sex in the past week, the intervention significantly increases reported enjoyment of sex. When overall sexual satisfaction is measured with a single global item asked of all women, the effect is positive but marginal, suggesting that coarser, less frequent measures may underestimate sexual well-being, and that intensive-margin measures should complement global ones to more comprehensively measure welfare, more broadly.

**Keywords:** family planning; experiment; unintended pregnancy; unwanted birth; Malawi

[a]**Corresponding Author**. Mailman School of Public Health, Columbia University. mvk2109@cumc.columbia.edu
[b]Mailman School of Public Health, Columbia University, New York, USA. sv2890@cumc.columbia.edu
[c]Harvard T.H. Chan School of Public Health, Harvard University, Boston, USA. dcanning@hsph.harvard.edu

# 1 Introduction

Family planning (FP) programs, and the field of sexual and reproductive health and rights (SRHR) that has been built around them, have long been oriented toward risk aversion, with efforts aimed at avoiding unintended pregnancies, unwanted births, and their downstream consequences for maternal and child health. However, the global agendas that motivate the SRHR field are explicitly framed in positive terms. Specifically, SRHR is invoked not merely as an absence of harm but as a right to be actively promoted, encompassing well-being and, in the World Health Organization (WHO)'s own working definition, "the possibility of having pleasurable and safe sexual experiences" (World Health Organization, 2006). The Guttmacher-Lancet Commission's integrated definition of SRHR makes the same affirmative claim at the level of global policy (Starrs et al., 2018). This leaves a persistent gap as to how the field articulates its goals and how it actually measures and operationalizes them, whereby programs are designed and assessed against negatively framed benchmarks, even as the frameworks that justify them promise something more affirmative.

Empirical evidence on how FP programs affect sexual health and well-being is limited and mixed, with most analyses to date being clinically oriented. For example, trials and observational studies have measured the impact of contraception on sexual function or dysfunction, or the side effects of specific contraceptive methods on desire, arousal, or satisfaction (e.g., Higgins and Smith (2016); Higgins et al. (2022)). This literature speaks to the physiological effects of particular contraceptive and reproductive technologies but says little about whether FP programs themselves, as comprehensive social and behavioral interventions, shape sexual enjoyment, pleasure, or satisfaction more broadly. A growing number of scholars have argued that pleasure and well-being deserve a more central place in the reproductive health research and measurement agenda, most notably Higgins and Hirsch's critique of the field's "pleasure deficit" (Higgins and Hirsch, 2007), in which they call to incorporate sexuality directly into research on contraceptive use (Higgins and Hirsch, 2008; Boydell et al., 2021), building on earlier arguments that pleasure-inclusive programming is an essential precondition for effective prevention, not a distraction from it (Philpott et al., 2006). From a well-being perspective, the imbalance is notable: the field has invested so heavily in identifying and averting bad outcomes (e.g., unintended pregnancy, unwanted births) that it has given comparatively little thought to what it would mean to systematically promote good ones (i.e., better sex).

Part of the reason for this gap is that sex remains a sensitive and stigmatized subject and is consequently under-measured in routine data collection. Demographic and Health Surveys (DHS) and similar instruments ask about the extensive margin (recent sexual activity)

but typically as a proxy for other indicators, such as pregnancy risk, or as a precursor for calibrating other indicators, like unmet need, rather than as an outcome in its own right. Even where the intensive margin is probed, it is usually limited to quantity, such as how many times a woman had sex in the past one or two weeks, and rarely quality (e.g., sexual pleasure, enjoyment, or satisfaction) (Gruskin et al., 2019; Mitchell et al., 2021). Where pleasure-inclusive measures have been fielded, early evidence suggests that they meaningfully change what interventions appear to accomplish (Zaneva et al., 2022), but such measures remain rare, and evaluations that assess the impact of contraception on these outcomes often frame them as secondary to physiological outcomes related to sexual (dys)function (Smith et al., 2014; Da et al., 2024). Moreover, little is known about whether FP interventions affect them through social and behavioral channels, whether the ability to plan or space a pregnancy shapes sexual satisfaction, as distinct from the clinical channels through which a given contraceptive method may affect sexual functioning.[1]

A related challenge is that even where sexual satisfaction is measured, the instruments that are currently available to measure it remain incompletely and unevenly validated. The Global Measure of Sexual Satisfaction (GMSEX), which is among the most widely used brief instruments in the literature, has only recently been validated simultaneously across multiple cultures, languages, genders, and sexual orientations (Paquette et al., 2025), and its underlying psychometric structure is still being actively interrogated using modern item response theory methods (Hay et al., 2026). Comparative work finds that single-item and multi-item satisfaction measures do not always behave interchangeably (Mark et al., 2014), and competing instruments, such as the New Sexual Satisfaction Scale, have themselves only been validated in a small number of contexts (tulhofer et al., 2010). Taken together, no single validated measure of sexual satisfaction or well-being yet exists for survey programs to adopt; the measurement framework itself is still being built even as global health agendas are asking programs to act on it.

This study contributes to both gaps using experimental data from a randomized controlled trial in urban Malawi. We field a richer battery of enjoyment and satisfaction measures alongside standard extensive- and intensive-margin sexual activity questions, and we leverage women's random assignment to an FP intervention to identify the causal effect of improved FP program access on their sexual well-being, rather than the effect of any method in isolation.

[1] A larger literature has examined the impact of contraception on sexual functioning from a clinical and physiological perspective; see, for example Casey et al. (2017); Li et al. (2004); Smith et al. (2014), among others.

## 2 Data and Experimental Design

We use longitudinal data from the Malawi Family Planning Study (MFPS), a randomized controlled trial that was conducted in Lilongwe, Malawi, between September 2016 and February 2019.[2] The MFPS evaluated a multicomponent FP intervention that addressed key barriers to contraceptive access, including knowledge gaps and constraints related to distance, service availability, and cost, among a cohort of married women who were pregnant or immediate postpartum at baseline. Women were interviewed annually over three survey waves.

Following a baseline survey, women were individually randomized into treatment or control arms. The treatment arm received: 1) up to six private, home-based FP counseling visits with trained counselors; 2) free, private transportation to a high-quality FP clinic; and 3) full coverage of all costs for FP services received at that clinic or elsewhere, including for methods, FP consultations, tests, and for the treatment for contraceptive-related side effects and contraindications. The control arm received publicly available information on FP methods and the availability of services at nearby clinics and was not recontacted until follow-up.

The MFPS has been the basis for several analyses of the intervention's effects (Karra et al., 2022; Maggio et al., 2024): two years after exposure, treated women were 5.9 percentage points (p.p.) more likely to use any contraceptive method, which was driven by increased adoption of long-acting reversible methods (particularly the implant), and 43 percent less likely to have a subsequent pregnancy within 24 months of their index birth. Pechar et al. (2026) further show that the intervention reduced within-woman variability in fertility preferences over time (without changing desired family size), particularly among women who were pregnant at baseline, and this stability was associated with higher contraceptive use. Montas et al. (2026) leverage random counselor assignment to show that higher-quality counseling reduces the number of sessions that intervention women receive while increasing the amount of information delivered within each session. Contraceptive use also increases among women who are exposed to higher-quality counselors. Finally, Karra et al. (2026) find that the intervention increased women's employment and hours worked after two years of exposure.

[2] A detailed protocol describing the study design and intervention can be found in Karra and Canning (2020).

## 2.1 Eligibility, Sample Size and Randomization

Eligible women were married; aged 18 to 35; permanent residents of Lilongwe; had not been sterilized or undergone a hysterectomy; and either currently pregnant or had given birth (immediate postpartum) within six months from the time of recruitment at baseline. If multiple eligible women were present in a household, the youngest woman was selected to participate. Enrolled women resided at least five households apart to limit spillovers. A total of 2,143 women were enrolled in 2016 and were randomized using stratified covariate-balanced randomization on parity, current contraceptive use, age at sexual debut, education, work status, and neighborhood, resulting in 1,027 women being assigned to the intervention arm and 1,113 women assigned to the control arm.[3] Table A1 shows the two arms are balanced on observable baseline characteristics, with only a minor difference in prior contraceptive use; a joint significance test similarly supports no systematic baseline differences across arms at baseline. Additional detail on the randomization protocol and covariate balance is presented in Karra and Canning (2020).

Of the 2,143 women enrolled, 1,672 (78 percent) were re-interviewed at endline. Maggio et al. (2024) and Pechar et al. (2026) find attrited and retained women to be statistically indistinguishable by treatment arm, supporting the internal validity of the experimental comparison, although attrition is related to some baseline characteristics. Specifically, women who were lost to follow-up were, on average, younger, less educated, of lower parity, and somewhat more likely to want additional children. We therefore interpret our estimates as representative of the retained sample.

## 2.2 Key Outcomes

Our primary outcomes capture sexual well-being along the extensive and intensive margins of sexual activity, drawn from a dedicated sexual health and marital satisfaction module fielded at the second (2017) and third (2018) survey waves.

On the extensive margin, we construct two measures of recency from women's reports of when they last had sex: 1) the natural log of days since her last sexual encounter; and 2) an indicator for whether she had sex within the past week. Both of these indicators are constructed for all women, regardless of recent sexual activity.

On the intensive margin, we restrict our analysis to women who report having had sex in the past week, and we measure both quantity and quality of sex within that window. Quantity is measured as the log number of times a woman had sex in the preceding week.

[3]Three women withdrew from the study following completion of the baseline and prior to randomization, resulting in a total of 2,140 women who were randomized into intervention and control arms.

Quality is captured by two items: 1) a woman's enjoyment of sex in the past week, measured on a four-point scale from "never" to "fully" enjoyed; and 2) a woman's satisfaction with her own sexual functioning over the same period, on a five-point scale. We additionally measure desire (self-reported frequency of sexual thoughts in the past week) on a five-point scale that, like the recency measures, is asked of all women rather than conditional on recent activity. We reverse-code each ordinal item so that higher values indicate greater desire, enjoyment, or satisfaction, and construct a composite sexual well-being index as the standardized (z-scored) average of the desire, enjoyment, functioning, and global satisfaction items among women who answered all four. Our final measure is women's overall satisfaction with their sex life, which is constructed on a 10-point scale and is asked of all women irrespective of recent sexual activity. Unlike the past-week items, this global measure is not conditioned on the extensive margin.

## 2.3 Covariates

The covariate-adjusted specification (model 2 in each table) conditions on nine baseline characteristics: neighborhood, age group, age at sexual debut, number of children alive, ever use of FP, education (primary or lower vs. secondary or higher), work status, religion (Christian vs. other), and ethnicity (Chewa vs. other). These are the same characteristics reported in Table A1, with one difference: the balance table reports age continuously for balance checking, while the regressions use a categorical 5-year age group version of the variable. We supplement this standard set with two covariates that are specific to our sexual well-being outcomes. First, we control for the number of days elapsed since a woman's index birth at the time of each survey wave, since time since birth is likely to be associated with the resumption and frequency of sexual activity, independent of treatment status. Second, we control for a woman's baseline sexual activity, measured as the natural log of days since her last sexual encounter at the time of the baseline survey. We include this as a covariate, rather than restricting the analytic sample to women who report recent sexual activity at each follow-up wave, because sexual activity at follow-up is itself plausibly affected by our intervention; conditioning the sample on a post-treatment characteristic would therefore risk introducing selection bias.

# 3 Empirical Strategy

Our main results present an Intent-to-Treat (ITT) analysis, identifying the average treatment effect of assignment to treatment relative to control:

$$Y_{ia} = \alpha + \beta_T T_{ia} + \mathbf{X}_{ia}^{0}\gamma + \delta_a + \varepsilon_{ia} \tag{1}$$

$Y_{ia}$ is the outcome of interest for woman $i$ in enumeration area $a$ (measures of sexual well-being); $T_{ia}$ indicates treatment assignment, with $\beta_T$ our parameter of interest; $\mathbf{X}_{ia}^{0}$ is a vector of baseline covariates, including randomization strata; $\delta_a$ is a neighborhood fixed effect; and $\varepsilon_{ia}$ is the error term. We report $\beta_T$ from a naïve model (treatment only) and from the fully adjusted model above, with heteroskedasticity-robust standard errors throughout.

# 4 Results

## 4.1 Sample and Baseline Balance

The analytic sample comprises 2,139 women enrolled at baseline; 1,113 women in the control arm and 1,026 women in the intervention arm. Table A1 compares baseline characteristics across arms; all normalized differences fall below the conventional 0.10 imbalance threshold (Austin, 2009), with ever use of FP at baseline showing the largest standardized difference ($-0.091$; 73.6% control vs. 77.5% treatment). The table's joint F-test ($p = 0.660$) indicates overall balance, and we find no basis for treating any single characteristic as a threat to randomization. With this said, we nevertheless retain the full covariate set in the adjusted specification as a means to reduce the residual variance and minimize our standard errors in our estimated treatment effects.

## 4.2 Recency of Sexual Activity: The Extensive Margin

Table A2 reports the treatment effect on how recently women report having had sex, both without (model 1) and with (model 2) baseline covariates. In 2017, women assigned to the FP intervention report having had sex more recently than control women; the log days since last sex measure falls by $-0.141$ (SE 0.068, $p = 0.038$) with covariates against a control mean of 1.878 (Panel A), and the probability of having had sex within the past week rises by 3.9 percentage points (p.p.) (SE 0.020, $p = 0.055$) against a control mean of 77.2 percent (Panel B), a 5.1 percent relative increase. Both effects are directionally consistent and similarly sized across the unadjusted and adjusted specifications. By 2018, neither effect persists,

and both coefficients are far from significant ($p = 0.590$ and $p = 0.577$, respectively). This pattern is not likely attributable to confounding given the random intervention assignment and that covariate adjustment shifts every estimate only marginally.

## 4.3 Quantity of Sex: The Intensive Margin

Table A3 restricts our analysis to women who report having had sex in the past week and asks whether the FP intervention changes coital frequency among these women. We find null effects across both years, where both the linear (Panel A) and Poisson (Panel B) specifications return coefficients indistinguishable from zero. In 2017, the intervention raises weekly coital frequency by an estimated 2.6 percent (Poisson IRR 1.026, $p = 0.576$) against a control mean of 2.9 times per week, and in 2018 the point estimate is negative and equally small (IRR 0.988, $p = 0.739$). This null result matters for interpreting Table A4 below: any intervention effect that may be observed on the quality of sex among those women who are already sexually active cannot be explained by the intervention simply inducing more sex among these women.

## 4.4 Quality of Sex: The Intensive and Extensive Margins

Table A4 reports the effects of the FP intervention on three measures of sexual quality: enjoyment and satisfaction with one's own sexual functioning, which are both asked of women who had sex in the past week (Panels A and B), and overall sexual satisfaction, which is asked of all women irrespective of recent sexual activity (Panel C). In 2017, women assigned to the FP intervention report significantly greater enjoyment of sex, a 0.126 point increase (SE 0.059, $p = 0.034$) against a control mean of 3.194 on the four-point scale, equivalent to a 3.9 percent relative increase; satisfaction with one's own sexual functioning shows a similarly sized but marginally significant effect (0.130, SE 0.073, $p = 0.074$) against a control mean of 3.924. Overall sexual satisfaction moves in the same direction (0.166, SE 0.146) against a control mean of 7.804 on the ten-point scale, but is not significant at conventional levels ($p = 0.253$). As with recency, none of these effects persist into 2018, where all three coefficients are small and statistically indistinguishable from zero.

The pattern across panels is notable, with the two measures that are conditioned on recent sexual activity (enjoyment and satisfaction with one's own sexual functioning) showing the clearest evidence of an intervention effect, while the single global item of sexual satisfaction that is asked of all women showing the same sign in 2017 but a generally weaker signal, and a null or even slightly opposite sign in 2018. We return to this distinction and its implications for how sexual well-being is measured in FP research in Section 5.

# 5 Discussion

In this study, we evaluate the causal impact of a FP intervention on women's sexual well-being using experimental evidence from Malawi. In 2017, all five sexual well-being outcomes that we measure (recency, having had sex in the past week, enjoyment, satisfaction with one's own sexual functioning, and global sexual satisfaction) move in the same positive direction for women who were assigned to the FP intervention. We find that a narrow measure of sexual enjoyment that is conditional on recent sexual activity detects this effect at conventional statistical significance levels; on the other hand, a measure that si based on a single global sexual satisfaction item detects an effect of comparable relative magnitude, once each outcome is scaled to its own response range, but fails to clear conventional significance ($p = 0.253$), driven by a standard error large enough to swamp a point estimate of otherwise comparable size. This finding is not naturally explained by the global measure being a coarser or less sensitive construct; in fact, a coarser measure would more likely attenuate the point estimate itself and not just inflate its standard error while simultaneously leaving the coefficient the largest of the five. It is, however, what reduced statistical power would likely produce - a real, directionally consistent effect that is measured with more noise.

This noise plausibly has a specific source. To our knowledge, this is among the first studies to field a global sexual satisfaction measure as an outcome of a FP field experiment, and the instruments underlying such measures, including GMSEX, have only recently been validated across contexts (Paquette et al., 2025). The psychometric literature on sexual satisfaction is still maturing even in well-resourced settings, and validation work in the low- and middle-income contexts where most FP investments actually occur remains sparser still. A first priority, then, is not simply to field more sexual well-being items, but rather to validate and adapt them for the populations in which FP programs operate, which in turn would allow researchers and program evaluators to be confident that a given instrument is measuring the same underlying construct across the settings that are being compared. An instrument that is deployed for the first time in a population for which it has not been validated, whose response scale may be unfamiliar, and whose translations have not been psychometrically tested would be expected to introduce a certain level of measurement error, and that error would manifest precisely as reduced precision on a real treatment effect rather than as a change in the effect's sign or its presence in the other, better-established measures. This is what we observe in this study, where the ceiling-heaping in the raw distribution of the global satisfaction measure, which constitutes a disproportionate share of responses clustering at the scale's maximum, is one concrete, observable symptom of an underlying scale measurement problem.

This finding reframes, rather than resolves, the measurement question that was raised earlier in this study. Specifically, the task ahead is not simply to argue that global measures are inherently too blunt to detect real effects (our own results caution against that conclusion), but rather to validate these instruments in the populations where FP programs actually operate. Doing so would allow future studies to distinguish a genuinely null effect from an underpowered true one without having to infer the difference indirectly from assessing results across several imperfectly comparable outcomes (as we do here).

A second, more practical challenge is infrastructural. Even a well-validated instrument does little for the field if it remains confined to a contextually specific trial. SRHR surveys that are conducted at scale, such as the DHS, already ask about recent sexual activity, but as a proxy for pregnancy risk rather than as a well-being outcome in its own right. Incorporating validated items that are oriented towards well-being into this existing infrastructure, alongside the clinical integration efforts that are already underway in contraceptive counseling settings (Garrett and Vaisey, 2022), would allow for sexual well-being to be tracked over time and across populations at a scale that no single field experiment could operationalize. Embedding a small number of well-validated items into existing instruments is a substantially lower bar to achieve than building new data collection systems from scratch, particularly given recent changes to global data infrastructure following the termination of USAID and the DHS program (Abdel Ghany et al., 2026).

The deeper implication, however, is not just methodological but evaluative. If sexual well-being were only of academic interest, better measurement of the construct would be a technical footnote. It is not. Individuals organize substantial parts of their relationships, decisions, and behavior around sex (Diamond and Huebner, 2012); it is not a peripheral domain of life but a central one, and FP is one of the few areas of public health policy that is positioned to act directly on it. Global institutions have already said as much (World Health Organization, 2006; Starrs et al., 2018; Grinde, 2022), and the World Association for Sexual Health has issued a formal declaration on sexual pleasure as a matter of health policy (Ford et al., 2021). What seems to be missing is not the mandate, but the follow-through. FP programs are still evaluated almost exclusively against negatively framed benchmarks, including contraceptive uptake, (dis)continuation, and averted pregnancy, while the affirmative outcomes these same institutions call for are unmeasured and, therefore, unrewarded. Early evidence suggests this gap is not merely rhetorical; pleasure-inclusive programming appears to add measurable value where it has been implemented (Zaneva et al., 2022). If FP programs are going to be evaluated on whether they are a success or a failure, sexual well-being belongs among the outcomes that valuation should be based on, not held apart from it as a secondary concern that is to be addressed only after the negatively framed FP

and reproductive health aims have been met.

# 6 Conclusion

Based on our findings, the extent to which this FP intervention improved women's sexual well-being depends on the margin by which well-being is being measured. On the extensive margin, women in the intervention arm report having had sex more recently, an effect that fades over the study period. On the intensive margin, the FP intervention does not increase how often women who are already sexually have sex, but it does increase how much they enjoy it, and (at reduced precision) how satisfied they are with their own sexual functioning and sex lives overall. These effects on the intensive margin are also visible only once the outcome is measured with enough granularity to detect them.

Beyond this study, our analysis contributes one of the few causal estimates of FP on a positively framed outcome of sexual well-being, situated against a much larger literature that is organized around the outcomes that FP programs seek to avert. It suggests that the field's difficulty in detecting these effects may reflect the inadequacy of available measurement tools as much as the absence of a true effect. A targeted, granular measure and a single global item pointed in the same direction here, but only the former did so with enough precision to be legible as a potential finding. If sexual well-being is to be treated as a program outcome in practice, and not only in the policy language that nominally signals it as a goal, the field will need to develop validated instruments that are embedded in surveys that researchers and program evaluators already use, together with a willingness to evaluate programmatic success in part by how people feel about their sex lives, and not only by whether they avoided an unintended outcome in having it.

# 7 Acknowledgments

This research uses original data collected by Canning and Karra with support from Innovations for Poverty Action (IPA) in Malawi. Karra would like to acknowledge the dedication and support of Carly Farver, Patrick Baxter, Bagrey Ngwira, Reginald Chunda, Viola Nyirongo, Violet Chitsulo, Macdonald Salamu, and the entire MFPS team. The findings, interpretations, and conclusions expressed in this study are entirely those of the authors.

## 7.1 Funding Sources

This project was supported by two grants from the Hewlett Foundation and the Program for Women's Empowerment Research (POWER) (Grant No: 2018-8083, 2020-1162).

## 7.2 Ethical Considerations and Trial Registration

The MFPS trial was registered at the American Economics Association Registry for randomized controlled trials on May 7, 2015 (AEARCTR-0000697) and at the Registry for International Development Impact Evaluations (RIDIE) on May 28, 2015 (RIDIE-STUDY-ID-556784ed86956). Ethical approval to conduct the study was received from the Harvard University IRB (protocol number IRB16-0421) and the Malawi National Health Sciences Research Committee (protocol number 16/7/1628).

# References


Abdel Ghany, J., A. A. Nur, K. L. D. MacQuarrie, J. Wilde, E. A. Sully, M. Karra, U. Gazeley, B. M. John, and L. Montana (2026, March). In the wake of USAID cuts, we can create a Demographic and Health Survey Program founded on more equitable data infrastructure and stronger research integrity. *Proceedings of the National Academy of Sciences 123*(9), e2513242123. 9

Austin, P. C. (2009, November). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine 28*(25), 3083–3107. 6

Boydell, V., K. Q. Wright, and R. D. Smith (2021, September). A Rapid Review of Sexual Pleasure in First Sexual Experience(s). *The Journal of Sex Research 58*(7), 850–862. 1

Casey, P. M., K. L. MacLaughlin, and S. S. Faubion (2017, March). Impact of Contraception on Female Sexual Function. *Journal of Women's Health 26*(3), 207–213. 2

Diamond, L. M. and D. M. Huebner (2012). Is Good Sex Good for You? Rethinking Sexuality and Health. *Social and Personality Psychology Compass 6*(1), 54–69. 9

Da, C. A., E. C. H. Moreira, and C. L. Zamboti (2024, August). Sexual function, quality of life, anxiety, and depression in women of reproductive age using hormonal, nonhormonal, and no contraceptive methods. *The Journal of Sexual Medicine 21*(8), 683–690. 2

Ford, J. V., E. Corona-Vargas, M. Cruz, J. D. Fortenberry, E. Kismodi, A. Philpott, E. Rubio-Aurioles, and E. Coleman (2021, October). The World Association for Sexual Healths Declaration on Sexual Pleasure: A Technical Guide. *International Journal of Sexual Health 33*(4), 612–642. 9

Garrett, C. C. and A. Vaisey (2022, February). Most women really actually do just appreciate being asked: clinicians views on integrating sexual wellbeing into contraceptive care. *Culture, Health & Sexuality 24*(3), 315–329. 9

Grinde, B. (2022, February). The Contribution of Sex to Quality of Life in Modern Societies. *Applied Research in Quality of Life 17*(1), 449–465. 9

Gruskin, S., V. Yadav, A. Castellanos-Usigli, G. Khizanishvili, and E. Kismdi (2019, January). Sexual health, sexual rights and sexual pleasure: meaningfully engaging the perfect triangle. *Sexual and Reproductive Health Matters 27*(1), 29–40. 2

Hay, A., C. F. Belu, and P. Santos-Iglesias (2026, May). Unpacking the Global Measure of Sexual Satisfaction: An Item Response Theory Approach. *The Journal of Sex Research 0*(0), 1–15. 2

Higgins, J. A. and J. S. Hirsch (2007). The Pleasure Deficit: Revisiting the "Sexuality Connection" in Reproductive Health. *Perspectives on Sexual and Reproductive Health 39*(4),

240–247. 1

Higgins, J. A. and J. S. Hirsch (2008, October). Pleasure, Power, and Inequality: Incorporating Sexuality Into Research on Contraceptive Use. *American Journal of Public Health 98*(10), 1803–1813. 1

Higgins, J. A., R. D. Kramer, K. Q. Wright, B. Everett, D. K. Turok, and J. N. Sanders (2022, May). Sexual Functioning, Satisfaction, and Well-Being Among Contraceptive Users: A Three-Month Assessment From the HER Salt Lake Contraceptive Initiative. *The Journal of Sex Research 59*(4), 435–444. 1

Higgins, J. A. and N. K. Smith (2016, May). The Sexual Acceptability of Contraception: Reviewing the Literature and Building a New Concept. *The Journal of Sex Research 53*(4-5), 417–456. 1

Karra, M. and D. Canning (2020). The Effect of Improved Access to Family Planning on Postpartum Women: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols 9*(8), e16697. 3, 4

Karra, M., D. Maggio, and D. Canning (2026). Family Planning Women's Employment, and Household Labor Supply. *Working Paper*. 3

Karra, M., D. Maggio, M. Guo, B. Ngwira, and D. Canning (2022). The causal effect of a family planning intervention on womens contraceptive use and birth spacing. *Proceedings of the National Academy of Sciences 119*(22), e2200279119. 3

Li, R. H. W., S. S. T. Lo, D. K. G. Teh, N.-C. Tong, M. H. Y. Tsui, K.-B. Cheung, and T. K. H. Chung (2004, December). Impact of common contraceptive methods on quality of life and sexual function in Hong Kong Chinese women. *Contraception 70*(6), 474–482. 2

Maggio, D., M. Karra, and D. Canning (2024, October). Family Planning and Children's Human Capital: Experimental Evidence From Urban Malawi. *Demography 61*(5), 1667–1698. 3, 4

Mark, K. P., D. Herbenick, J. D. Fortenberry, S. Sanders, and M. Reece (2014, February). A Psychometric Comparison of Three Scales and a Single-Item Measure to Assess Sexual Satisfaction. *The Journal of Sex Research 51*(2), 159–169. 2

Mitchell, K. R., R. Lewis, L. F. O'Sullivan, and J. D. Fortenberry (2021, August). What is sexual wellbeing and why does it matter for public health? *The Lancet Public Health 6*(8), e608–e613. 2

Montas, M. C., S. Vincent, and M. Karra (2026). The Causal Effect of Counseling Quality: Experimental Evidence from Urban Malawi. 3

Paquette, M.-M., S. Bergeron, N. Bigras, M. Kos, L. Nagy, S. W. Kraus, Z. Demetrovics,

M. N. Potenza, R. Ballester-Arnal, D. Batthyny, J. Billieux, P. Briken, J. Burkauskas, G. Crdenas-Lpez, J. Carvalho, J. Castro-Calvo, L. Chen, G. Ciocca, O. Corazza, R. I. Csako, D. P. Fernandez, E. F. Fernandez, H. Fujiwara, J. Fuss, R. Gabrhelk, A. Gewirtz-Meydan, B. Gjoneska, M. Gola, J. B. Grubbs, H. T. Hashim, M. S. Islam, M. Ismail, M. C. Jimnez-Martnez, T. Jurin, O. Kalina, V. Klein, A. Klt, S.-K. Lee, K. Lewczuk, C.-Y. Lin, C. Lochner, S. Lpez-Alvarado, K. Lukavsk, P. Mayta-Tristn, D. J. Miller, O. Orosov, G. Orosz, H. Chang, K. Park, F. P. Ponce, G. R. Quintana, G. C. Quintero Garzola, J. Ramos-Diaz, K. Rigaud, A. Rousseau, M. D. T. Scanavino, M. K. Schulmeyer, P. Sharan, M. Shibata, S. Shoib, V. Sigre-Leirs, L. Sniewski, O. Spasovski, V. Steibliene, D. J. Stein, C. Strong, B. C. nsal, M.-P. Vaillancourt-Morel, M. C. Van Hout, and B. Bthe (2025). Sexual satisfaction across cultures, genders, languages, and sexual orientations: Validation of the Global Measure of Sexual Satisfaction. *Psychology of Sexual Orientation and Gender Diversity*. 2, 8

Pechar, S., D. Maggio, and M. Karra (2026). Family Planning and the Dynamics of Fertility Preferences: Experimental Evidence from Urban Malawi. 3, 4

Philpott, A., W. Knerr, and V. Boydell (2006, January). Pleasure and Prevention: When Good Sex Is Safer Sex. *Reproductive Health Matters 14*(28), 23–31. 1

Smith, N. K., K. N. Jozkowski, and S. A. Sanders (2014, February). Hormonal Contraception and Female Pain, Orgasm and Sexual Pleasure. *The Journal of Sexual Medicine 11*(2), 462–470. 2

Starrs, A. M., A. C. Ezeh, G. Barker, A. Basu, J. T. Bertrand, R. Blum, A. M. Coll-Seck, A. Grover, L. Laski, M. Roa, Z. A. Sathar, L. Say, G. I. Serour, S. Singh, K. Stenberg, M. Temmerman, A. Biddlecom, A. Popinchalk, C. Summers, and L. S. Ashford (2018, June). Accelerate progresssexual and reproductive health and rights for all: report of the GuttmacherLancet Commission. *The Lancet 391*(10140), 2642–2692. 1, 9

World Health Organization (2006). Defining Sexual Health: Report of a Technical Consultation on Sexual Health, 28-31 January 2002. Technical report, World Health Organization, Geneva, Switzerland. 1, 9

Zaneva, M., A. Philpott, A. Singh, G. Larsson, and L. Gonsalves (2022, February). What is the added value of incorporating pleasure in sexual health interventions? A systematic review and meta-analysis. *PLOS ONE 17*(2), e0261034. 2, 9

tulhofer, A., V. Buko, and P. Brouillard (2010, July). Development and Bicultural Validation of the New Sexual Satisfaction Scale. *The Journal of Sex Research 47*(4), 257–268. 2

# A Appendix

Figure A1: Histogram of Sexual Satisfaction (Global 10-Point Scale, by Arm

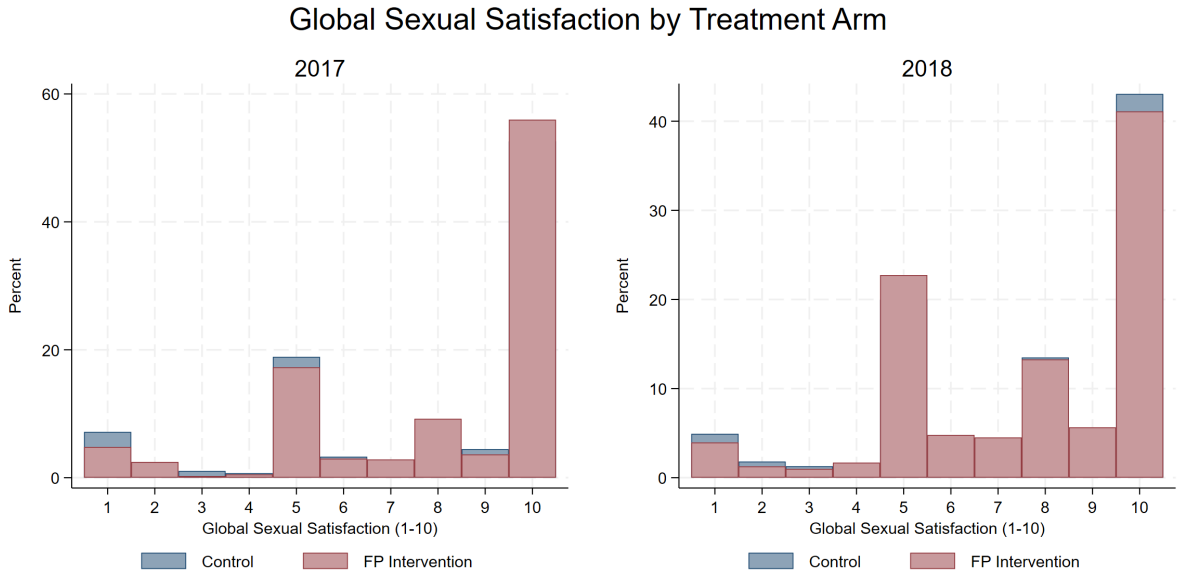

Table A1: Balance Table, Baseline Covariates

| | **Estimation Sample** | | **Control Group** | | **Treatment Group** | | |
|---|---|---|---|---|---|---|---|
| | N | Mean | N | Mean | N | Mean | Normalized Difference |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Area | 2139 | 43.588 | 1113 | 43.749 | 1026 | 43.413 | 0.019 |
| Age Group | 2139 | 1.726 | 1113 | 1.734 | 1026 | 1.717 | 0.017 |
| Age at First Sex (Years) | 2130 | 18.857 | 1107 | 18.810 | 1023 | 18.907 | -0.035 |
| Total number of children alive | 2138 | 2.255 | 1112 | 2.226 | 1026 | 2.287 | -0.049 |
| Ever Use of FP (1 = Yes) | 2139 | 0.755 | 1113 | 0.736 | 1026 | 0.775 | -0.091 |
| Primary Education (1 = Yes) | 2139 | 0.587 | 1113 | 0.588 | 1026 | 0.586 | 0.004 |
| Worked in Last Week (1 = Yes) | 2139 | 0.096 | 1113 | 0.093 | 1026 | 0.099 | -0.023 |
| Religion (1 = Christian) | 2139 | 0.827 | 1113 | 0.816 | 1026 | 0.838 | -0.059 |
| Ethnicity (1 = Chewa) | 2137 | 0.418 | 1112 | 0.421 | 1025 | 0.415 | 0.013 |
| Log(1 + Days Since Last Sex) | 2123 | 2.284 | 1103 | 2.292 | 1020 | 2.275 | 0.012 |
| Hormonal Method Use (1 = Yes) | 2139 | 0.218 | 1113 | 0.217 | 1026 | 0.220 | -0.009 |
| Had Sex in Past Week at Baseline (1 = Yes) | 2123 | 0.625 | 1103 | 0.622 | 1020 | 0.628 | -0.013 |
| Observations for F-test | | | | | | | 2121 |
| F-test of joint significance: p-value | | | | | | | 0.735 |

*Note:* Columns (1)-(2) pool both arms; Columns (3)-(4) are the control arm; Columns (5)-(6) are the treatment arm; Column (7) is the normalized difference $(\bar{x}_T - \bar{x}_C)/\sqrt{(s_T^2 + s_C^2)/2}$. The reported $F$-test is a test of joint significance.

Table A2: Main ITT Effects: Recency of Sexual Activity (Extensive Margin)

| **Panel: A** | **Log(1 + Days Since Last Sex)** | | | |
|---|---|---|---|---|
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | -0.167** | -0.141** | 0.044 | 0.041 |
| | (0.066) | (0.068) | (0.074) | (0.077) |
| | [0.012] | [0.038] | [0.549] | [0.590] |
| Observations | 1653 | 1511 | 1515 | 1365 |
| $R^2$ | 0.004 | 0.063 | 0.000 | 0.039 |
| Control Mean | 1.878 | 1.878 | 1.870 | 1.870 |
| **Panel: B** | **Had Sex in the Past Week (0/1)** | | | |
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | 0.044** | 0.039* | -0.015 | -0.013 |
| | (0.020) | (0.020) | (0.022) | (0.023) |
| | [0.026] | [0.055] | [0.504] | [0.577] |
| Observations | 1653 | 1511 | 1515 | 1365 |
| $R^2$ | 0.003 | 0.047 | 0.000 | 0.025 |
| Control Mean | 0.772 | 0.772 | 0.775 | 0.775 |
| Controls | | x | | x |

*Note:* Each column reports coefficients from a separate regression of the row outcome on FP intervention assignment; model (1) includes no controls, model (2) includes the full covariate set. Panel A measures recency of last sexual activity as ln(1+days since last sex), where a negative coefficient indicates more recent activity; Panel B is an indicator for whether the respondent had sex within the past week. Both outcomes are asked of all women regardless of recent sexual activity, unlike the past-week-conditional measures in Tables 3 and 4. Controls include area, age group, age at sexual debut, number of children alive, ever use of FP, education, work status, religion, ethnicity, time since index birth (exposure), and baseline (2016) sexual activity. Robust standard errors in parentheses; p-values in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Main ITT Effects: Quantity of Sexual Activity (Intensive Margin)

| **Panel: A** | **Log(1 + Times Had Sex, Past Week) – OLS** | | | |
|---|---|---|---|---|
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | 0.027 | 0.012 | -0.000 | -0.005 |
| | (0.026) | (0.026) | (0.025) | (0.025) |
| | [0.295] | [0.659] | [0.987] | [0.851] |
| Observations | 1250 | 1149 | 1111 | 1005 |
| $R^2$ | 0.001 | 0.036 | 0.000 | 0.049 |
| Control Mean | 1.263 | 1.263 | 1.268 | 1.268 |
| **Panel: B** | **Times Had Sex, Past Week – Poisson (IRR)** | | | |
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| Times had sex in the past week | | | | |
| FP Intervention | 1.063 | 1.026 | 0.995 | 0.988 |
| | (0.049) | (0.048) | (0.034) | (0.035) |
| | [0.182] | [0.576] | [0.888] | [0.739] |
| Observations | 1250 | 1149 | 1111 | 1005 |
| Control Mean (raw count) | 2.942 | 2.942 | 2.874 | 2.874 |
| Controls | | x | | x |

*Note:* Each column reports coefficients from a separate regression of the row outcome on FP intervention assignment, among women who report having had sex in the past week; model (1) includes no controls, model (2) includes the full covariate set. Panel A estimates a linear model of ln(1+times had sex in the past week) via OLS. Panel B estimates a Poisson model of the raw count of times had sex in the past week; reported coefficients are incidence rate ratios (IRR), where a value of 1 indicates no treatment effect. Unlike Table 2, both panels condition on having had sex in the past week, since the frequency question is only asked of women who report recent sexual activity. Controls include area, age group, age at sexual debut, number of children alive, ever use of FP, education, work status, religion, ethnicity, time since index birth (exposure), and baseline (2016) sexual activity. Robust standard errors in parentheses; p-values in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Main ITT Effects: Quality of Sexual Activity (Intensive and Extensive Margins)

| **Panel: A** | **Enjoyment/Pleasure, Past Week** | | | |
|---|---|---|---|---|
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | 0.096* | 0.126** | 0.029 | 0.026 |
| | (0.057) | (0.059) | (0.058) | (0.062) |
| | [0.090] | [0.034] | [0.618] | [0.677] |
| Observations | 1282 | 1179 | 1138 | 1031 |
| $R^2$ | 0.002 | 0.035 | 0.000 | 0.019 |
| Control Mean | 3.194 | 3.194 | 3.185 | 3.185 |
| **Panel: B** | **Own Sexual Functioning Satisfaction, Past Week** | | | |
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | 0.117* | 0.130* | 0.012 | 0.002 |
| | (0.070) | (0.073) | (0.070) | (0.074) |
| | [0.093] | [0.074] | [0.861] | [0.980] |
| Observations | 1288 | 1184 | 1142 | 1035 |
| $R^2$ | 0.002 | 0.029 | 0.000 | 0.020 |
| Control Mean | 3.924 | 3.924 | 3.975 | 3.975 |
| **Panel: C** | **Global Sexual Satisfaction** | | | |
| | 2017 | | 2018 | |
| | (1) | (2) | (1) | (2) |
| FP Intervention | 0.227 | 0.166 | -0.029 | -0.099 |
| | (0.138) | (0.146) | (0.138) | (0.147) |
| | [0.101] | [0.253] | [0.835] | [0.500] |
| Observations | 1622 | 1485 | 1479 | 1335 |
| $R^2$ | 0.002 | 0.023 | 0.000 | 0.013 |
| Control Mean | 7.804 | 7.804 | 7.613 | 7.613 |
| Controls | | x | | x |

*Note:* Each column reports coefficients from a separate regression of the row outcome on FP intervention assignment; model (1) includes no controls, model (2) includes the full covariate set. Panels A and B measure enjoyment of sex (1-4 scale) and satisfaction with one's own sexual functioning (1-5 scale) in the past week, respectively; both are reverse-coded so higher values indicate more positive experiences, and, as in Table 3, both condition on having had sex in the past week. Panel C measures overall sexual satisfaction on a 1-10 scale and, unlike Panels A and B, is asked of all women regardless of recent sexual activity. Controls include area, age group, age at sexual debut, number of children alive, ever use of FP, education, work status, religion, ethnicity, time since index birth (exposure), and baseline (2016) sexual activity. Robust standard errors in parentheses; p-values in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.