Family Planning and Women's Labor Supply*

Mahesh Karra† Daniel Maggio‡ David Canning§

December 2, 2024

Abstract

We conduct a randomized controlled trial that provided pregnant and postpartum women with comprehensive access to family planning services. We assess the causal impact of the family planning intervention on women's work and labor market outcomes. We find that women are 5.1 percentage points more likely to be working for a wage or are self-employed after two years of intervention exposure, which is driven by a 3.4 percentage point increase in wage-earning labor. Our results are primarily driven by women who were pregnant at baseline, experienced fewer short inter-pregnancy interval births, and participated more in household decision-making.

*This 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). This research makes use of original data collected by Canning and Karra with support from Innovations for Poverty Action (IPA) in Malawi. The authors 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 Malawi Family Planning Study (MFPS) team, which comprised of 22 enumerators and 7 counselors over a three year study period. This project was supported by two grants from the William and Flora Hewlett Foundation and the Human Capital Initiative's Program for Women's Empowerment Research (POWER) at the Boston University Global Development Policy Center. Ethical approval to conduct the study was received from the Harvard University Institutional Review Board (protocol number IRB16-0421) and from the Malawi National Health Sciences Research Committee (protocol number 16/7/1628). Informed consent was obtained from all participants in the trial. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of Innovation for Poverty Action and its affiliated organizations, or those of the Executive Directors of the Innovation for Poverty Action they represent.

†Frederick S. Pardee School of Global Studies, Boston University, Boston, MA USA. mvkarra@bu.edu.

‡Department of Economics, Rutgers University, New Brunswick, NJ USA. daniel.maggio@rutgers.edu.

§Harvard T.H. Chan School of Public Health, Boston, MA USA. dcanning@hsph.harvard.edu.

1 Introduction

Expanded access to family planning (FP) is often proposed as an effective development strategy. From a rights-based perspective, FP promotes women's autonomy and freedom of choice through increased agency to shape their own reproductive lives (UNFPA, 2014; Newman and Feldman-Jacobs, 2015; Hardee and Jordan, 2021). As a means to many ends, FP is also considered to be of instrumental importance given its significant and positive contributions to a broader range of outcomes, including fertility, health and human capital, women's employment and work, and overall social and economic well-being (Canning and Schultz, 2012; Miller, 2010; Joshi and Schultz, 2007). These links have been theorized as part of the broader macroeconomic evidence base on the Demographic Dividend, where access to FP and reproductive health services plays a key role in enabling women and couples to realize their fertility preferences by averting unintended pregnancies and mistimed births, contributing to positive fertility change and reduced uncertainty over birth events (Bloom et al., 2003, 2007; Lee and Mason, 2006; Becker and Tomes, 1976; Becker and Lewis, 1973; Cavalcanti et al., 2021). These impacts on fertility, in turn, affect human capital as women are better able to invest in their health and education, as well as in the health and education of their children. Moreover, if labor markets are positioned to absorb additional labor supply, FP may also lead to increases in women's employment and labor force participation as their time and resource constraints are relaxed. These potential impacts on labor supply can allow women to economically contribute more, which would contribute to overall economic growth at the aggregate level and improve the welfare of women and their families at the household level (Bloom et al., 2003).

There is a large quasi-experimental literature that has investigated the causal relationship between fertility and women's labor market outcomes, leveraging variation in non-singleton births (Rosenzweig and Wolpin, 1980b), sex-selective childbearing (Angrist and Evans, 1996; Chun and Oh, 2002; Cruces and Galiani, 2007), and fertility restrictions (Li and Zhang, 2007; Huang et al., 2021) for identification. Most of these findings suggest that fertility decline has large and positive effects on women's labor supply. Meanwhile, much of the high-quality evidence linking contraceptive access and FP to women's work comes from high-income contexts. Perhaps the most influential works from this literature study the introduction of oral contraceptive methods ("the pill") on women's labor supply and show that the availability of oral contraceptives facilitated the attainment of professional education among women (Goldin and Katz, 2002), lowered their likelihood of pregnancy early in their lives (Bailey, 2006), and decreased the wage gap between men and women (Bailey et al., 2012). However, evidence on the link between FP and women's work in low-income settings is much more limited and mixed. In these settings, the strongest evidence to date comes from the Matlab FP, Maternal and Child Health Program (FP-MCH) in Bangladesh. Schultz (2009) show that two decades after the initiation of the FP-MCH program, women in program areas earned 40 percent more than women in control areas, in part as a result of a 17 percent reduction in fertility. While this evidence is promising, it warrants three caveats. First, a recent long-run analysis by Barham et al. (2021) suggests that this effect has dissipated 35 years after the intervention's initial rollout. Second, the Matlab FP-MCH program combined FP services with child health services, specifically a wide-scale vaccination campaign, making it difficult to isolate the labor market effects of FP from other components that were simultaneously rolled out. Finally, we note that the majority of the literature has presumed that the impact of FP on downstream labor market outcomes is, to a large extent, mediated through fertility decline. However, in contexts where preferences for birth spacing dominate preferences for limiting childbearing, we may observe changes to women's labor supply that are driven by changes in access to and use of FP services even when total parity over a woman's reproductive lifetime does not change.

In this study, we estimate the causal effect of improved access to FP services on women's labor market outcomes by means of a randomized controlled trial, which was conducted from 2016 to 2019 in urban Malawi. As part of the trial, pregnant or immediate postpartum women were recruited and were randomized to either a control arm or an intervention arm following a baseline survey. Women assigned to the intervention arm received a package of services that consisted of up to six free FP counseling sessions, free transport to a local high-quality FP clinic, and financial reimbursement for FP commodities and related services for a two-year period. Two annual follow-up surveys were conducted following the baseline and rollout of the intervention.

While results on the primary outcomes from the trial are reported in Karra et al. (2022), this study reports intent-to-treat (ITT) estimates on downstream outcomes related to women's work after two years of exposure to our intervention. We show that women who received the comprehensive FP package were 5.1 percentage points (p.p.) more likely to be working for a wage or were self-employed two years following the intervention. These changes seem to be largely driven by an increase in wage-earning work, where we show that women in the intervention arm are 3.4 p.p. more likely to report working for labor income compared to women in the control arm. To examine the possibility of intra-household substitution in labor supply, we also examine the impacts of the intervention on husbands' labor market outcomes. For men, we do not observe evidence of an effect on overall waged work or self-employment, unemployment, or labor force participation throughout the study. We do, however, show that husbands are 5.4 p.p. less likely to report being self-employed at endline and 6.3 p.p. more likely to report working for a wage, providing some evidence that they substitute between types of labor. These results on both women's and men's labor market outcomes suggest that the intervention increased the overall labor supply in the household.

Our paper makes two key contributions. First, we are able to identify the causal effects of family planning. While there is extensive literature showing the positive effects of contraceptive access on women's work in the United States (Goldin and Katz, 2002; Bailey, 2006; Bailey et al., 2012), this literature is less developed in low-income settings. Literature in this space has, to date, provided a range of evidence to highlight the positive impacts of FP on first-stage outcomes such as contraceptive use and birth spacing (Bhatia et al., 1980; Schultz, 2009; Joshi and Schultz, 2007; Debpuur et al., 2002). However, there is less high-quality evidence linking FP to more downstream outcomes related to women's social and economic well-being, and evidence of the economic impacts of FP is mixed. On the one hand, studies by Miller (2010) and Joshi and Schultz (2007) have shown positive impacts of FP on women's labor market outcomes. In contrast, Barham et al. (2021) found that women exposed to the Matlab FP-MCH program experienced no significant improvements in their economic well-being, although the authors noted the challenge to disentangling the impact of FP from the role of other maternal and child health services that were also offered as part of the program. Our study contributes to this ongoing discussion by providing more direct experimental evidence of the causal impact of a family planning-focused intervention.

Second, through our experimental design, we are able to study the causal links between FP and women's labor market outcomes, where the evidence to date has primarily relied on observational or quasi-experimental studies to estimate the extent to which reductions in fertility and access to contraceptives contribute to women's work and labor supply (refer to Angrist and Evans (1996) and Goldin and Katz (2002) as examples). We add to this evidence base by showing that the results we observe are driven by reductions in short inter-pregnancy interval births and improvements in decision-making power among women who were pregnant at baseline.

The rest of this paper proceeds as follows. Section 2 provides the theoretical framework that links FP to economic outcomes. Section 3 describes the intervention and experimental design. Section 4 outlines our empirical strategy, and Section 5 presents the results from our analysis of the intervention's impact and explores potential mechanisms. We discuss the implications of our findings and conclude in Section 6. We present several secondary analyses in the Appendix.

2 Theoretical Framework

We expect our FP intervention to impact women's labor market outcomes in three ways. First, improvements in birth spacing through our intervention may relax constraints that are placed on women as a result of a pregnancy, which in turn may allow women to return to and more effectively engage in the labor market (Bloom et al., 2009; Rosenzweig and Wolpin, 1980a). We confirm in Karra et al. (2022) that improved access to FP through our intervention was associated with a 44 percent reduction in the relative risk of short inter-pregnancy intervals (a pregnancy within two years of a preceding birth), implying that constraints associated with fertility and postpartum childrearing could be significantly reduced for women in our intervention arm. Second, even among women who would not have become pregnant without improved access to FP services, our intervention could enable these women and couples to more effectively control their fertility and reduce uncertainty surrounding pregnancy and birth (Cavalcanti et al., 2021). This improved control may, in turn, encourage women to engage in the labor market by enabling them to devote more time and resources to activities beyond childcare with greater certainty. Finally, increased contraceptive use, and the use of hormonal contraception, may also confer health benefits through reductions in the risk of maternal anemia and maternal nutritional depletion that may follow pregnancy and childbirth (Gebremedhin and Asefa, 2018; Bellizzi and Ali, 2018; Karra and Canning, 2024). These effects from contraceptive use may, in turn, contribute to improvements in women's labor market outcomes through improved health and human capital attainment.

While there is likely a significant correlation between fertility and women's labor market outcomes, better labor market conditions may, in turn, cause reductions in fertility as higher income increases the opportunity cost of childbearing (Becker, 1960; Schultz, 1969; Rosenzweig and Wolpin, 1980a; Ager et al., 2020). In response to this increased opportunity cost, women and couples may reduce their birth parity over their reproductive lifetimes. Since women in our sample have not yet reached this life stage, we cannot directly estimate this relationship in this study; however, we are able to show correlational evidence that short inter-pregnancy intervals significantly reduce the likelihood of women working and that these types of pregnancies decrease as a result of our intervention.

3 Study Design

Our study leverages a randomized controlled trial (RCT) that was conducted in Lilongwe, Malawi, from November 2016 to February 2019. We provide an abbreviated description of the trial in relation to our present analysis; a more detailed description of the study protocol and the implementation of the intervention is presented elsewhere Karra and Canning (2020).

As part of the trial, 2,143 women[1] who were either pregnant or immediately postpartum (had given birth within 6 months of the baseline screening), between the ages of 18 and 35, and living in Lilongwe were recruited in 2016. Following a baseline survey, women were individually randomized to either an intervention arm or a control arm. Women assigned to the intervention arm received a comprehensive FP package of services that consisted of private, home-based FP counseling sessions, free transport to a high-quality FP clinic in Lilongwe, and financial reimbursement for all contraceptive methods and related services over a two year period. In Karra et al. (2022), we show that treatment assignment is largely balanced across a range of observable baseline characteristics. In Table 1, we supplement these findings with a balance table of women's labor market characteristics at baseline, and we again do not observe any large or statistically significant differences across treatment arms.[2] In our sample, 23.5 percent of women reported either working for a wage or being self-employed, at baseline, with a majority (20.1 percent) of these women reported being self-employed. We find the opposite for men, 92.4 percent of whom report engaging in either activity, with a majority (64.4 percent) engaging in wage-earning labor. We also observe much higher labor force participation rates among men than women, which partly may signal gender disparities in the labor market but also reflects the fact that half of the sample women are pregnant and may, therefore, not be working at baseline.

The data used in this study are from three in-person surveys that were conducted annually throughout the course of the trial. A baseline survey was implemented from September 2016 to January 2017, and two follow-up surveys were conducted with women on an annual basis. Data collection for the first follow-up survey (midline) began in August 2017 and was completed in February 2018, and data collection for the second follow-up survey (endline) began in August 2018 and was completed in February 2019. In Karra et al. (2022), we show that 76 percent of the recruited sample was successfully followed up at endline, and we do not find evidence of differential loss to follow-up across treatment arms. To minimize attrition, abbreviated phone surveys were conducted at each follow-up survey with any women who had either moved out of Lilongwe or who were unable to be reached in person after three attempts. However, not all secondary outcomes that we present in this study were collected as part of the phone survey; as a result, the analytic samples in this study are smaller than

[1]Three women are later dropped from the sample because they withdrew consent following completion of the baseline survey.

[2]The full balance tables for the trial are presented in Karra et al. (2022) and Maggio et al. (2024). For this study, we have included a balance table of the variables that we use as covariates in our adjusted ITT specifications and those variables that are used to explore mechanisms in the Appendix.

those presented in Karra et al. (2022).[3]

3.1 The Intervention

Women assigned to the intervention arm were offered a comprehensive, multi-component FP package over a two-year service period. Designed in coordination with the Malawi Ministry of Health's Reproductive Health Directorate (RHD), local health authorities, NGOs working in FP and reproductive health, and local communities, the intervention aimed to simultaneously overcome multiple access barriers that were identified through formative fieldwork and piloting. As such, the package combined FP counseling with free transport and financial reimbursement to overcome key barriers to FP care-seeking in Lilongwe. By offering a comprehensive package, this study aims to determine the causal effect of improved access to FP.

As part of the intervention, women were offered up to six free, private FP counseling sessions that were conducted by trained counselors at their homes over a two-year period. During each session, women received detailed information on a full range of contraceptive methods as well as information related to the terms and conditions of their use. Women also received brochures that presented information on the benefits of healthy timing and spacing of births, the benefits of breastfeeding, the potential benefits of FP on maternal and child health, and the management of contraceptive-related side effects and contraindications. Importantly, women were also informed of the other intervention components during the first counseling session.

In addition, women assigned to the intervention arm were offered free transportation (private taxi rides) to a local private FP clinic, the Good Health Kauma Clinic, which offered women comprehensive FP services and care.[4] Transport was provided by a driver who was hired exclusively for the intervention. In addition to the driver, women who sought to utilize the transport service were accompanied by a female field manager in the taxi, which served to mitigate any social stigma surrounding a woman traveling alone in the company of another man.

Finally, women assigned to the intervention arm received up to 17,500 MWK (∼$25.00 USD) in financial coverage for any costs that they incurred for receiving FP care at the Kauma Clinic or any other clinic in Lilongwe that women chose to attend. Costs that

[3]Using less detailed questions, we are able to measure some of the labor market outcomes for those women who were surveyed by phone. Specifically, we can measure women's work; but we cannot infer women's labor force participation or unemployment or labor market outcomes for their husbands.

[4]Services offered by the Kauma Clinic include the insertion and removal of long-acting methods, referrals for sterilization, pregnancy tests, FP for women with HIV or other health concerns, and comprehensive treatment and management of side effects.

were eligible to be reimbursed included those related to the procurement and removal of contraceptive methods, FP consultations, costs related to the treatment for and management of contraceptive-related side effects and contraindications, lab tests and fees, and other exam fees related to FP and reproductive health care. Though the reimbursement allowance amount was non-transferable to non-FP-related services, women were allowed to redeem any reimbursement amount over multiple visits throughout the two-year intervention period.

Women assigned to the control arm received publicly available information on contraceptive methods and information about their nearest FP clinic. The women were given this information during the baseline survey and were only contacted again during annual follow-up surveys.

4 Empirical Strategy

4.1 ITT Effects

In this study, we present results from ITT specifications of the treatment effect of our intervention on outcomes related to women's and men's work, self-employment, and labor force participation after two years of exposure to our intervention.[5] For our ITT estimates, we run the following standard specification:

$$Y_{it} = \alpha + \beta_T T_i + \gamma Y_{i0} + \varepsilon_{it} \tag{1}$$

where Y_{it} is the value of the dependent variable for woman i at time t. In this study, we use several dependent variables related to the extensive and intensive margins of women's labor market outcomes and time use, which we discuss in more detail below. In our regressions, T_i represents the treatment assignment for women i, Y_{i0} represents the baseline value of the outcome, and ε_{it} represents the heteroskedastic-robust error term.

We also estimate the following adjusted ITT specifications:

$$Y_{it} = \alpha + \beta_T T_i + \gamma Y_{i0} + \mathbf{X}_{i0}\zeta + \mathbf{Z}_{i0}\eta + \varepsilon_{it} \tag{2}$$

where $\mathbf{X}_{i0}$ represents the vector of covariates that were used to balance assignment to treatment at baseline, and $\mathbf{Z}_{i0}$ represents the vector of other baseline covariates that are adjusted for in our regressions.

In our models, the coefficient β_T identifies the intent-to-treat effect of being assigned to

[5] In the Appendix, we report the treatment effect on secondary measures of work status (agricultural labor and unemployment), ITT results after one year of intervention exposure, and results on women's time use).

the treatment arm on outcomes. We choose to study the effect of treatment assignment rather than treatment uptake (i.e adherence-adjusted treatment effects) in this analysis for two reasons. The first is a function of how we implemented our intervention. Specifically, women who were assigned to the intervention arm were only introduced to all three FP intervention components during their first FP visit by the counselor; as a result, all women assigned to the intervention arm were, in some sense, "treated," and it is therefore unclear how we would define treatment compliance otherwise. Secondly, studying the intent-to-treat effect allows us to infer the population-level effect of improved access to FP services in a context where service uptake may be only relevant and endogenously demanded by specific subgroups of women and not demanded by others (e.g. women who seek to become pregnant again soon or who are trying to conceive), conditional on being offered.

In this study, we present results on a range of outcomes related to women's labor supply. Our main outcome of interest is women's work, where we specifically focus on women's wage-earning work or self-employment within the past week. We report results for both men and women for each outcome of interest. In the Appendix, we also report results on the impact of our intervention on agricultural labor supply, unemployment, and time use.

4.2 Sample Attrition

As mentioned in Section 3, some of the women included in our sample were interviewed via phone survey, resulting in smaller samples than those reported in Karra et al. (2022). This is particularly true of men's labor outcomes, which were not pre-registered in our study and thus were excluded from our phone survey modules due to time and resource constraints. At endline, were able to re-interview 1,672 of our 2,143 sample women, yielding a 78 percent follow-up rate. In contrast, we only obtained employment information for 1,392 husbands, a 65 percent follow-up rate.

In Maggio et al. (2024), we show that there is similar attrition rate among children enrolled in our sample and provide a detailed analysis of this sample attrition. This analysis shows that our sample attrition is non-differential in levels and characteristics across intervention arms. However, regardless of the intervention arm, attrition is correlated with a woman's education and religion. We conduct a similar attrition analysis, which is presented in the Appendix, by providing a balance table of attrition across men's and women's labor market outcomes. Similarly, we see no evidence of differential attrition invalidating our initial randomization. However, women lost to attrition were 7-8 p.p. less likely to work for a wage or be self-employed at baseline.

Although the concerns about attrition are not as severe in this analysis as those presented in Maggio et al. (2024), the high level of sample attrition among men suggests that it may be worthwhile to explore how this attrition affects our estimates. We do this in the following two ways.

4.2.1 Inverse Propensity Score Weighting

In following Macours and Molina Millán (2017), we first correct for sample attrition using a propensity score weighting approach. This approach allows us to adjust for the imbalances in sample attrition that are correlated with women's education, religion, and employment discussed above. We accompany our main ITT specification with estimates from an inverse probability weighting adjustment that uses the estimated propensity scores as weights.

4.2.2 Bounding

We supplement these propensity score weighted estimates by placing bounds on our estimates in the Kling-Leibman-Katz tradition Kling et al. (2007). This bounding analysis assumes that the missing outcome data are a given number of standard deviations away from the within-intervention arm mean. We follow the methodology used in Maggio et al. (2024) and calculate the bounds under two sets of assumptions, whereby the missing data takes the outcome within 0.1 and 0.2 standard deviations of the group mean, respectively.

5 Results

5.1 ITT Effects

We first describe our main outcomes over time. Figure 1 presents the likelihood of working over the study period for both men and women. For women, we observe similar trends among the treatment and control arms of our intervention during the first year of the intervention.[6] However, during the second year of the intervention, we see that the likelihood of working among women in the intervention arm increases at a faster rate than among women in the control arm.

In Table 2, we investigate these trends empirically and present the ITT results on women's labor market outcomes at the second-year follow-up. In confirming the trend observed in Figure 1, we observe that women assigned to the intervention arm of our study were 5.1 p.p. more likely to work for a wage or be self-employed. This effect seems to be driven by increased participation in wage-earning labor; women assigned to the intervention arm are

[6]The monotonic improvement in outcomes across groups throughout the intervention period is likely reflective of women non-differentially reentering the labor market following their pregnancy.

3.4 p.p. more likely to report being engaged in wage-earning labor at endline. While our point estimates on women's work are positive, our estimates on labor force participation are imprecise. With this said, we find that unemployment over both waves is close to zero, suggesting that any women who entered the labor market were likely to become employed.

Regarding our attrition adjustments, we see that, across outcomes, our propensity-weighted estimates closely match our main adjusted intent-to-treat estimate. This would suggest that the imbalance we observe between attritors and non-attritors may not significantly impede our ability to make externally valid claims about the effect of our intervention.

From the bounds on our estimates, we see that our estimates quickly lose significance and switch signs for one outcome (self-employment) under the more conservative assumptions. While these results suggest that our study may not have sufficient statistical power to detect smaller effects than our estimates suggest for our intervention, they also show that the differential attrition would have be quite severe to invalidate our results on women's labor outcomes.

Our findings are promising for understanding the broader welfare implications of FP. However, we may also expect to observe unintended negative consequences if households view female and male labor supply as substitutes; if this is the case, then we may observe a commensurate reduction in male labor market outcomes towards female labor market outcomes as fertility falls. To explore this possibility, we document trends in male work and labor supply in Figure 1. In contrast to women's work, we observe a lack of trends between wage-earning work and self-employment in men. The insignificant differences in work between treatment arms provide some evidence of substitution away from wage-earning labor towards self-employment among men in the control arm, while the composition of labor activity among men in the intervention arm remains relatively constant. We further evaluate these relationships in our empirical analysis of the intervention on men's labor outcomes in Table 3. Overall, we do not observe a compensatory decrease in men's labor outcomes in response to increases in women's work, suggesting an increase in the total household labor supply as a result of the intervention. As suggested by Figure 1, we find a 5.2 p.p. difference in men's likelihood of self-employment between treatment arms. Conversely, we find evidence of a compensating difference in wage-earning labor, although this estimate loses significance when controlling for differences in the outcome at baseline. These results indicate that as FP services allow women to increase their participation in the labor market, their husbands dedicate less of their time toward household enterprises, opting instead to participate in more formal, waged types of work.

Given the higher level of attrition that we observe, the estimates in Table 3 are more sensitive to our bounding analysis. However, in the cases where we observe statistically significant results in self-employment and wage-earning work, we see that the assumptions made on our missing outcome data must be fairly severe, and we would argue overly conservative, to reverse the direction of our treatment effects.

5.2 Potential Mechanisms

In Section 2, we describe how we expect FP to affect labor force outcomes through changes in fertility, birth spacing, and control over birth timing. To explore those mechanisms, one would typically look to instruments for fertility using our randomly assigned treatment assignment, or turn to conducting a causal mediation analysis. In our setting, neither of these approaches is likely to provide convincing evidence. Regarding the former, although we show in Karra et al. (2022) that our intervention causes a 44 percent relative decrease in short-term fertility, its explanatory power over this fertility decline (which would be the first stage in an instrumental variables regression) is low. Regarding the latter, as we expect work and fertility decisions to be highly correlated and confounded, we are unlikely to find a covariate set that allows us to satisfy the sequential ignorability assumption Imai et al. (2010).

Instead, we decompose the traditional mediation analysis and do not make strong causal claims when discussing potential mechanisms. In doing so, we present the treatment effect of our intervention on two potentially mediating variables. The first is a dummy variable that takes the value one if a women had a subsequent pregnancy following the index pregnancy that made her eligible for enrollment into the study. The second is a summary-weighted index of women's relative decision-making power.[7] In the Appendix, we show that the ITT results we observe are heterogeneous by a woman's pregnancy status at baseline, and so we decompose this analysis by that dimension to understand how pregnancy status at baseline may moderate any potential causal pathways that we observe.

We present this analysis in Table 4. Although we do not claim the findings in Panel C to be causal, this decomposed analysis is helpful in understanding the mechanisms that underlie our results. To confirm the findings in the Appendix, we observe a 9.2 percentage point increase in wage-earning work and self-employment and an 8.2 percentage point increase in labor force participation among the subgroup of women who were randomly assigned to the intervention and who were pregnant at baseline relative to women who were pregnant at baseline but assigned to the control arm. Panels B and C provide suggestive evidence that two reinforcing trends drive these heterogeneous effects. First, although the reduction in

[7] We calculate the index over dummy variables that take on a value of 1 if a woman reports that she has full control or shares control with her partner regarding decisions over her healthcare, major household purchases, minor household purchases, and visits to relatives.

subsequent pregnancies is similar across pregnant and non-pregnant women (3 p.p. versus 5 p.p.), these pregnancies are much more deleterious for pregnant women at baseline. A pregnant woman who experiences a subsequent pregnancy during our two-year follow-up period is 27 p.p. less likely to be working than her peers who did not have a subsequent pregnancy. This is substantially larger than the negative association for non-pregnant women (women who were immediate postpartum at baseline), who are only 15 percentage points less likely to be working if they had a subsequent pregnancy. This heterogeneity is likely driven by the fact that to have a subsequent pregnancy within our study period, pregnant women would necessarily have a short inter-pregnancy interval, which would likely affect their economic prospects more than women who spaced their births over a longer period. Second, we see that although the correlation between women's empowerment and work is strong for both pregnant and non-pregnant women, we only observe treatment effects on empowerment for women who were pregnant at baseline. This may relate to the timing of when households make decisions about healthcare spending on children and that the intervention, when provided to pregnant women, was able to preempt those decisions.

6 Discussion

In this study, we leverage experimental evidence from urban Malawi to identify the labor supply effects of improved access to family planning and reproductive health services. Our intent-to-treat estimates show large labor market gains for women who received improved access to FP; these women are 5.1 p.p. (15 percent of the endline control mean) more likely to work and 3.4 p.p. (29 percent) more likely to participate in wage-earning labor after two years of exposure to the intervention. In the Appendix, we show that these gains are especially large for women who were pregnant when they were introduced to the intervention; these women are 8.1 p.p. more likely to participate in the labor market. Furthermore, we do not observe deleterious effects on men's labor market outcomes, suggesting minimal, if any, compensating decreases in male labor supply due to the intervention.

Our results contribute to a large but mixed evidence base that has aimed to link women's reproductive behavior to their labor market outcomes (Angrist and Evans, 1996; Chun and Oh, 2002; Cruces and Galiani, 2007; Bailey et al., 2012). We also contribute to a growing literature on the long-term impacts and effectiveness of FP services, some of which has recently come into question. Our results stand in contrast to the findings of Barham et al. (2021) who do not find evidence of labor market improvements among women who received FP services. Methodologically, we supplement a predominantly observational and quasi-experimental literature with causal evidence from a randomized controlled trial.

Much like our prior experimental findings (Karra et al., 2022; Maggio et al., 2024), the results that we present are a part of a more extensive body of evidence that makes the case for improving access to high-quality FP services to all who demand them. All governments and organizations have budget constraints. If cost-benefit analyses of FP services were to only account for the positive effects of investing in FP on first-order outcomes (e.g. fertility, aversion of unintended pregnancy, health) while ignoring the more downstream positive externalities of these investments, they are likely to under-invest in these services, which have been found to be cost-effective (Kohler and Karra, 2019; Kohler, 2012). To this end, our findings suggest that policymakers must consider the effect of family planning on outcomes beyond contraceptive use, fertility, and even health when determining the merits, and the true impacts, of such programs.

References

Ager, P., B. Herz, and M. Brueckner (2020, October). Structural Change and the Fertility Transition. *The Review of Economics and Statistics 102*(4), 806–822. 4

Anderson, M. L. (2008, December). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495. Publisher: Taylor & Francis. 24

Angrist, J. D. and W. N. Evans (1996). Children and their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. NBER Working Paper No. 5778, National Bureau of Economic Research, Cambridge, MA. 1, 3, 12

Bailey, M. J. (2006, February). More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply*. *The Quarterly Journal of Economics 121*(1), 289–320. 1, 3

Bailey, M. J., B. Hershbein, and A. R. Miller (2012, July). The Opt-In Revolution? Contraception and the Gender Gap in Wages. *American Economic Journal: Applied Economics 4*(3), 225–254. 1, 3, 12

Barham, T., B. Champion, A. D. Foster, J. D. Hamadani, W. C. Jochem, G. Kagy, R. Kuhn, J. Menken, A. Razzaque, E. D. Root, and P. S. Turner (2021, July). Thirty-five years later: Long-term effects of the Matlab maternal and child health/family planning program on older women's well-being. *Proceedings of the National Academy of Sciences 118*(28). Publisher: National Academy of Sciences Section: Social Sciences. 2, 3, 12

Becker, G. (1960). An Economic Analysis of Fertility. NBER Chapters, National Bureau of Economic Research, Inc. 4

Becker, G. S. and H. G. Lewis (1973, March). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy 81*(2, Part 2), S279–S288. Publisher: The University of Chicago Press. 1

Becker, G. S. and N. Tomes (1976, August). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy 84*(4, Part 2), S143–S162. tex.ids= becker_child_1976-4 publisher: The University of Chicago Press. 1

Bellizzi, S. and M. M. Ali (2018, March). Effect of oral contraception on anemia in 12 low- and middle-income countries. *Contraception 97*(3), 236–242. 4

Bhatia, S., W. H. Mosley, A. S. G. Faruque, and J. Chakraborty (1980). The Matlab Family Planning-Health Services Project. *Studies in Family Planning 11*(6), 202–212. 3

Bloom, D., D. Canning, G. Fink, and J. Finlay (2007, June). Realizing the Demographic Dividend: Is Africa any different? PGDA Working Paper, Program on the Global Demography of Aging. 1

Bloom, D., D. Canning, and J. Sevilla (2003, February). *The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*. Rand Corporation. Google-Books-ID: 36rNSRG4r7YC. 1

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay (2009, June). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth 14*(2), 79–101. 4

Canning, D. and T. P. Schultz (2012). The economic consequences of reproductive health and family planning. *Lancet 380*(9837), 165–171. 1

Cavalcanti, T., G. Kocharkov, and C. Santos (2021, February). Family Planning and Development: Aggregate Effects of Contraceptive Use. *The Economic Journal 131*(634), 624–657. 1, 4

Chun, H. and J. Oh (2002, August). An instrumental variable estimate of the effect of fertility on the labour force participation of married women. *Applied Economics Letters 9*(10), 631–634. 1, 12

Cruces, G. and S. Galiani (2007, June). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics 14*(3), 565–573. 1, 12

Debpuur, C., J. F. Phillips, E. F. Jackson, A. Nazzar, P. Ngom, and F. N. Binka (2002, June). The Impact of the Navrongo Project on Contraceptive Knowledge and Use, Reproductive Preferences, and Fertility. *Studies in Family Planning 33*(2), 141–164. 3

Gebremedhin, S. and A. Asefa (2018, November). Association between type of contraceptive use and haemoglobin status among women of reproductive age in 24 sub-Saharan Africa countries. *BMJ sexual & reproductive health*, bmjsrh–2018–200178. 4

Goldin, C. and L. F. Katz (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy 110*(4), 730–770. Publisher: The University of Chicago Press. 1, 3

Hardee, K. and S. Jordan (2021). Advancing Rights-Based Family Planning from 2020 to 2030. *Open Access Journal of Contraception 12*, 157–171. 1

Huang, W., X. Lei, and A. Sun (2021, September). Fertility Restrictions and Life Cycle Outcomes: Evidence from the One-Child Policy in China. *The Review of Economics and Statistics 103*(4), 694–710. 1

Imai, K., L. Keele, and T. Yamamoto (2010, February). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science 25*(1), 51–71. Publisher: Institute of Mathematical Statistics. 11

Joshi, S. and T. P. Schultz (2007, February). Family Planning as an Investment in Development: Evaluation of a Program's Consequences in Matlab, Bangladesh. SSRN Scholarly

Paper ID 962938, Social Science Research Network, Rochester, NY. 1, 3

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. Company: JMIR Research Protocols Distributor: JMIR Research Protocols Institution: JMIR Research Protocols Label: JMIR Research Protocols Publisher: JMIR Publications Inc., Toronto, Canada. 5

Karra, M. and D. Canning (2024). Hormonal Contraception and Hemoglobin Levels in Urban Malawi: A Longitudinal Analysis. *Working Paper*. 4

Karra, M., D. Maggio, M. Guo, B. Ngwira, and D. Canning (2022, May). The causal effect of a family planning intervention on women's contraceptive use and birth spacing. *Proceedings of the National Academy of Sciences 119*(22), e2200279119. Publisher: Proceedings of the National Academy of Sciences. 2, 4, 5, 6, 8, 11, 13, 24, 28

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2007.00733.x. 9

Kohler, H.-P. (2012). Copenhagen Consensus 2012: Challenge Paper on "Population Growth". pp. 79. 13

Kohler, H.-P. and M. Karra (2019). The Benefits and Costs of Scaling Up Family Planning in Africa. Best Buys for Africa, Copenhagen Consensus Center, Copenhagen, Denmark. 13

Lee, R. and A. Mason (2006, August). Back to Basics: What is the Demographic Dividend? *Finance & Development 0043*(003). ISBN: 9781451922486 Publisher: International Monetary Fund Section: Finance & Development. 1

Li, H. and J. Zhang (2007, February). Do High Birth Rates Hamper Economic Growth? *The Review of Economics and Statistics 89*(1), 110–117. 1

Macours, K. and T. Molina Millán (2017, April). Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias. 9

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. 5, 8, 9, 13, 24

Miller, G. (2010). Contraception as Development? New Evidence from Family Planning in Colombia. *The Economic Journal 120*(545), 709–736. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2009.02306.x. 1, 3

Newman, K. and C. Feldman-Jacobs (2015). Family Planning and Human Rights - What's the Connection and Why is it Important? Technical report, Population Reference Bureau, Washington D.C. 1

Rosenzweig, M. R. and K. I. Wolpin (1980a, April). Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy 88*(2), 328–348. Publisher: The University of Chicago Press. 4

Rosenzweig, M. R. and K. I. Wolpin (1980b). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica 48*(1), 227–240. tex.ids= rosenzweig_testing_1980-2 publisher: [Wiley, Econometric Society]. 1

Schultz, T. P. (1969, March). An Economic Model of Family Planning and Fertility. *Journal of Political Economy 77*(2), 153–180. 4

Schultz, T. P. (2009). How does Family Planning Promote Development? Evidence from a Social Experiment in Matlab, Bangladesh, 1977 – 1996. 2, 3

UNFPA (2014). Programme of Action, adopted at the International Conference on Population and Development. Technical report, United Nations Population Fund, Cairo, Egypt. 1

Table 1: Baseline Balance Table for Labor Market Characteristics

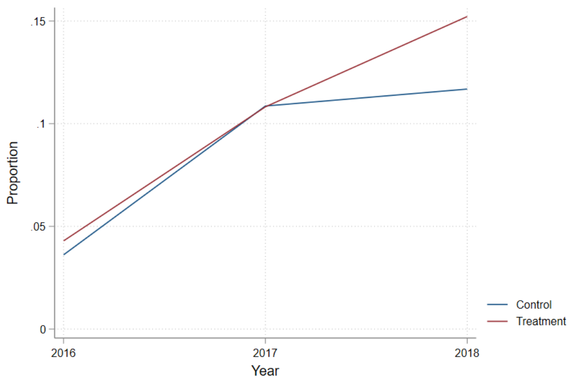
	(1) Control N=1,113	(2) Treatment N=1,027	(3) Total N=2,140	Difference (1)-(2)
Women's Labor Outcomes				
Woman Earns a Wage or is Self-Employed	0.232	0.237	0.235	-0.005
Wage Earning Work	0.036	0.043	0.039	-0.007
Self-Employed	0.204	0.197	0.201	0.007
Agricultural Work	0.005	0.009	0.007	-0.003
Woman is Unemployed	0.038	0.043	0.040	-0.005
Labor Force Participation	0.419	0.421	0.420	-0.002
Men's Labor Outcomes				
Man Earns a Wage or is Self-Employed	0.920	0.928	0.924	-0.008
Wage Earning Work	0.633	0.661	0.646	-0.028
Self-Employed	0.337	0.332	0.335	0.004
Agricultural Work	0.015	0.019	0.017	-0.003
Man is Unemployed	0.017	0.012	0.015	0.005
Labor Force Participation	0.989	0.991	0.990	-0.002
P-Value of F test for Joint Significance				**0.421**
Observations included in F-Test				**2124**

$*p < 0.1, **p < 0.05$

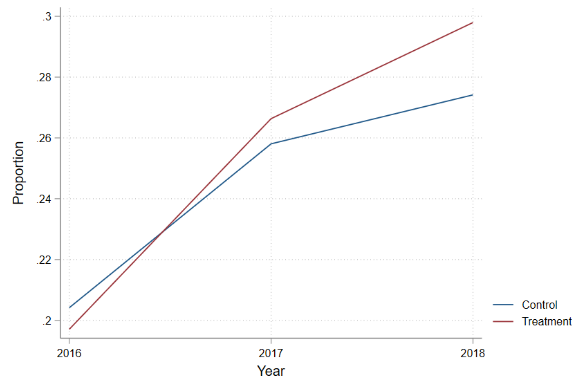
Notes: For all columns, the unit of observation is a household. Stars are based on the critical value from individual t-tests. Joint tests are computed using only observations for which we possess full data on all characteristics.

Figure 1: Work Outcome Trends over Time

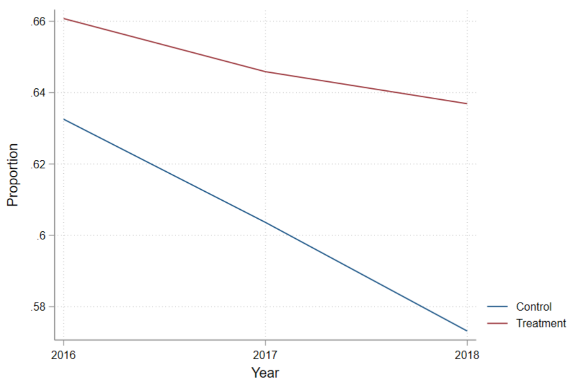
(a) Women's Wage-Earning Work



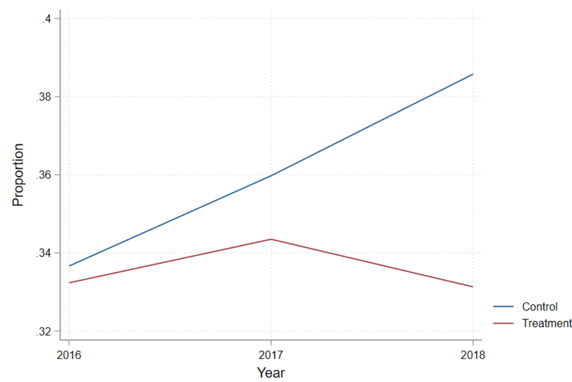
(b) Women's Self-Employment



(c) Men's Wage-Earning Work



(d) Men's Self-Employment



Notes: Lines represent the within-intervention arm outcome means at baseline, first-year follow-up and endline.

Table 2: ITT Effects on Women's Work, Year 2

			Bounds and Fully Adjusted Estimates					
	(1) Unadjusted	(2) ANCOVA	(3) +/- 0.2 SD	(4) +/- 0.1 SD	(5) Adjusted Estimate	(6) IPW	(7) +/- 0.1 SD	(8) +/- 0.2 SD
Panel A: Waged Work and Self-Employment								
Treatment	0.0515**	0.0512**	0.00625	0.0276	0.0479**	0.0511**	0.0703***	0.0916***
	(0.0241)	(0.0237)	(0.0186)	(0.0186)	(0.0237)	(0.0237)	(0.0186)	(0.0187)
Control Mean	0.382	0.382			0.382			
Panel B: Wage-Earning Work								
Treatment	0.0353**	0.0338**	0.00273	0.0176	0.0324*	0.0339**	0.0473***	0.0622***
	(0.0168)	(0.0166)	(0.0131)	(0.0131)	(0.0169)	(0.0168)	(0.0131)	(0.0132)
Control Mean	0.117	0.117			0.117			
Panel C: Self-Employment								
Treatment	0.0238	0.0265	-0.0144	0.00527	0.0247	0.0276	0.0446***	0.0642***
	(0.0222)	(0.0216)	(0.0171)	(0.0170)	(0.0217)	(0.0216)	(0.0170)	(0.0171)
Control Mean	0.274	0.274			0.274			
Panel D: Labor Force Participation								
Treatment	0.0378	0.0405	0.0204	0.0296	0.0369	0.0426*	0.0480**	0.0572**
	(0.0254)	(0.0249)	(0.0227)	(0.0226)	(0.0249)	(0.0251)	(0.0227)	(0.0227)
Control Mean	0.556	0.556			0.556			
Observations	1672	1668	2125	2125	1666	1666	2125	2125

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: For all columns, the unit of observation is a woman. The results presented are from OLS models with standard errors in parenthesis. OLS estimates are accompanied by estimates from an inverse probability weighting model (Column 6), and Kling-Leibman-type bounds (Columns 3, 4, 7, & 8). The adjusted regressions (Column 5) include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

Table 3: ITT Effects on Men's Work, Year 2

			Bounds and Fully Adjusted Estimates					
	(1) Unadjusted	(2) ANCOVA	(3) +/- 0.2 SD	(4) +/- 0.1 SD	(5) Adjusted Estimate	(6) IPW	(7) +/- 0.1 SD	(8) +/- 0.2 SD
Panel A: Waged Work and Self-Employment								
Treatment	0.0169	0.0161	-0.0203**	-0.00198	0.0148	0.0150	0.0346***	0.0528***
	(0.0141)	(0.0141)	(0.00940)	(0.00936)	(0.0142)	(0.0140)	(0.00936)	(0.00941)
Control Mean	0.917	0.917			0.917			
Panel B: Wage-Earning Work								
Treatment	0.0637**	0.0430*	-0.0133	0.0206	0.0397*	0.0434*	0.0884***	0.122***
	(0.0262)	(0.0237)	(0.0161)	(0.0160)	(0.0237)	(0.0240)	(0.0161)	(0.0162)
Control Mean	0.573	0.573			0.573			
Panel C: Self-Employment								
Treatment	-0.0545**	-0.0520**	-0.120***	-0.0863***	-0.0503**	-0.0516**	-0.0198	0.0135
	(0.0257)	(0.0223)	(0.0155)	(0.0154)	(0.0224)	(0.0227)	(0.0153)	(0.0154)
Control Mean	0.386	0.386			0.386			
Panel D: Labor Force Participation								
Treatment	0.0112	0.00962	0.00151	0.00590	0.00931	0.00877	0.0147***	0.0191***
	(0.00702)	(0.00681)	(0.00563)	(0.00562)	(0.00676)	(0.00673)	(0.00562)	(0.00564)
Control Mean	0.977	0.977			0.977			
Observations	1392	1384	2118	2118	1383	1383	2118	2118

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: For all columns, the unit of observation is a man. The results presented are from OLS models with standard errors in parenthesis. OLS estimates are accompanied by estimates from an inverse probability weighting model (Column 6), and Kling-Leibman-type bounds (Columns 3, 4, 7, & 8). The adjusted regressions (Column 5) include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

Table 4: Mechanism Analysis

Condition	Pregnant At Baseline				Not Pregnant At Baseline			
Outcome	Waged Work and Self-Employment		Labor Force Participation		Waged Work and Self-Employment		Labor Force Participation	
Mediator	Next Pregnancy	Empowerment	Next Pregnancy	Empowerment	Next Pregnancy	Empowerment	Next Pregnancy	Empowerment
Panel A: Treatment Effect on Outcome								
	0.0922***	0.0922***	0.0816**	0.0816**	0.00776	0.00776	-0.00138	-0.00138
	(0.0331)	(0.0331)	(0.0355)	(0.0355)	(0.0339)	(0.0339)	(0.0350)	(0.0350)
Panel B: Treatment Effect on Mediator								
	-0.0308*	0.148**	-0.0308*	0.148**	-0.0500**	-0.0778	-0.0500**	-0.0778
	(0.0161)	(0.0748)	(0.0161)	(0.0748)	(0.0198)	(0.0708)	(0.0198)	(0.0708)
Panel C: Correlation between Mediator and Outcome								
	-0.279***	0.0501***	-0.303***	0.0540***	-0.150***	0.0797***	-0.175***	0.0829***
	(0.0562)	(0.0167)	(0.0756)	(0.0181)	(0.0542)	(0.0172)	(0.0627)	(0.0184)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: For all columns, the unit of observation is a woman. Heteroskedasticity-robust standard errors are reported in parentheses.

A Appendix A: Variable Definitions

Outcomes:

1. **Wage-Earning Work or Self-Employment:** Indicator variable that takes a value of 1 if the respondent reports working for a wage, salary, commission or any payment in kind, or running a business of any kind for themselves or another household member within the last week, and 0 otherwise.

2. **Wage-Earning Work:** Indicator variable that takes 1 if the respondent reports working for a wage, salary, commission or any payment in kind within the last week, and 0 otherwise.

3. **Self-Employment:** Indicator variable that takes a value of 1 if the respondent reports running a business of any kind for themselves or another household member within the last week, and 0 otherwise.

4. **Unemployment:** Indicator variable that takes a value of 1 if the respondent reports seeking employment, seeking to start a business, or ”waiting for the busy season” (e.g. planting season in Malawi, when there is stronger labor demand), and 0 otherwise.

5. **Labor Force Participation:** Indicator variable that takes a value of 1 if the respondent reports any form of employment, including agricultural work, apprenticeships, or unpaid work, or report being unemployed as defined above, and 0 otherwise.

Covariates:

1. **Woman's Age**: The age of the respondent at the time of the survey, in years.

2. **Age at First Cohabitation**: Age at which the respondent and their spouse began living in the same household, in years.

3. **Number of Surviving Children**: Total number of *surviving* children the respondent had at the time of the survey.

4. **Ever Use of Contraception**: Indicator variable that takes a value of 1 if the respondent reported ever using contraception, and 0 otherwise.

5. **Primary Education**: Indicator variable that takes 1 if the respondent completed primary education, and 0 otherwise.

6. **Religion**: Indicator variable that takes a value of 1 if the respondent reported being Christian, and 0 otherwise.

7. **Ethnicity**: Indicator variable that takes a value of 1 if the respondent reported being of the Chewa tribe, and 0 otherwise.

8. **Contraceptive Use**: Indicator variable that takes a value of 1 if the respondent reports using contraception at baseline, and 0 otherwise.

9. **Empowerment Index**: A summary weighted index of indicator variables that take a value of 1 if a woman says she has full control or joint control over decisions regarding healthcare, major household purchases, minor household purchases, and visits to relatives, and 0 otherwise. The index is created by taking GLS-weighted averages of the component variables, following Anderson (2008) and using the Stata command *swindex*.

B Appendix B: Covariate Balance

Table B1 presents a balance table for the covariates used in our adjusted ITT estimates and the mediators used in our mechanism analysis. In our analyses, baseline values of our covariates are used, while endline values of the mediators are used; as such, the test of balance is presented at the relevant time point for each variable.

As we previously showed in Karra et al. (2022) and Maggio et al. (2024), our intervention arms are largely balanced across treatment arms at baseline, apart from a minor imbalance in women's ever use of contraception. Similarly, to confirm the findings that were reported in Karra et al. (2022), we note that women in the control arm are 4 percentage points more likely than women in the intervention arm to have had a subsequent pregnancy by the second-year endline.

C Appendix C: Attrition Balance

D Appendix D: Women's Time Use

In Section 2 of the main text, we outlined potential reasons why we may expect family planning services to improve women's labor market outcomes, citing the relaxation of constraints from childcare and other related responsibilities that women face. Within our data, we can test this hypothesis by presenting the effects of our intervention on women's time use.[8] In Table D1, we present these results for the first-year follow-up and the second-year endline.

[8]Time use is calculated using a recall diary of the women's primary activity for each of the past 24 hours from the time of interview.

Table B1: Balance Table for Covariates and Mediators

	(1) Control N=1,113	(2) Treatment N=1,027	(3) Total N=2,143	Difference (1)-(2)
Covariates (at Baseline)				
Woman's Age	24.509	24.657	24.580	-0.148
Age of First Cohabitation	18.810	18.904	18.855	-0.094
Number of Surviving Children	2.226	2.287	2.255	-0.061
Ever Use of Contraception	0.736	0.775	0.755	-0.039**
Primary Education	0.588	0.586	0.587	0.001
Any Work	0.093	0.099	0.096	-0.007
Religion (1 = Christian)	0.816	0.838	0.827	-0.023
Ethnicity (1 = Chewa)	0.421	0.414	0.418	0.007
Pregnancy Status	0.516	0.515	0.515	0.001
Contraceptive Use	0.235	0.239	0.237	-0.003
Mediators (at Endline)				
Subsequent Pregnancy by 2018	0.093	0.053	0.074	0.040***
Empowerment Index	-0.001	0.036	0.016	-0.036

*$p < 0.1$,**$p < 0.05$

Notes: For all columns, the unit of observation is a household. Stars are based on the critical value from individual t-tests.

Table C1: Attrition Balance Table

	Surveyed at Endline		Lost to Follow-Up					
	Control	Treatment	Control	Treatment	(1)-(2)	(1)-(3)	(2)-(4)	(3)-(4)
	(1)	(2)	(3)	(4)				
Panel A: Women's Outcomes								
Waged Work and Self-Employment	0.249	0.253	0.163	0.185	-0.004	0.087***	0.068**	-0.022
	[0.015]	[0.016]	[0.025]	[0.025]				
Wage-Earning W	0.041	0.045	0.018	0.037	-0.004	0.023**	0.008	-0.019
	[0.007]	[0.007]	[0.009]	[0.012]				
Self-Employment	0.219	0.212	0.145	0.148	0.007	0.074***	0.064**	-0.003
	[0.014]	[0.015]	[0.024]	[0.023]				
Labor Force Participation	0.445	0.428	0.317	0.396	0.016	0.128***	0.032	-0.079*
	[0.017]	[0.018]	[0.031]	[0.032]				
Observations	884	782	221	243				
Panel B: Men's Outcomes								
Waged Work and Self-Employment	0.919	0.938	0.924	0.911	-0.019	-0.005	0.026	0.012
	[0.010]	[0.009]	[0.014]	[0.015]				
Wage-Earning Labor	0.622	0.665	0.653	0.654	-0.043*	-0.031	0.011	-0.001
	[0.018]	[0.018]	[0.024]	[0.025]				
Self-Employment	0.348	0.346	0.316	0.307	0.002	0.032	0.038	0.008
	[0.018]	[0.019]	[0.024]	[0.024]				
Labor Force Participation	0.985	0.991	0.997	0.992	-0.006	-0.013**	-0.001	0.006
	[0.005]	[0.004]	[0.003]	[0.005]				
Observations	725	659	380	361				

*** p<0.01, ** p<0.05, * p<0.1

Notes: For all columns, the unit of observation is a woman in panel A and a man in panel B. Standard deviations are reported in brackets.

Table D1: ITT Effects on Women's Time Use

	(1) Work	(2) Childcare	(3) Housework	(4) Education
Panel A: First-Year Follow up				
ITT Effect	0.0581	0.0431	0.130	0.0303
	(0.139)	(0.0767)	(0.150)	(0.0217)
Observations	1657	1657	1657	1657
R Squared	0.0505	0.0234	0.102	0.00124
Panel B: Study Endline				
ITT Effects	0.249	-0.00771	-0.117	0.0413
	(0.165)	(0.0849)	(0.155)	(0.0299)
Observations	1515	1515	1515	1515
R Squared	0.0211	0.0292	0.0465	0.00137

Notes: For all columns, the unit of observation is a women. Time use is measured in number of hours during the 24-hour period prior to the interview. The results presented are from OLS models with standard errors in parenthesis. Results presented are obtained from an ITT estimation of the naïve treatment effect on outcomes, adjusted only by the value of the outcome at baseline. Heteroskedasticity-robust standard errors are reported.

Across waves, we do not observe evidence of an intervention effect on women's time use. At the study's endline, our estimates are quite imprecise, but the coefficient signs imply that women spend more time at work and substitute away from both housework and child care. Given the observed effects on labor outcomes, we interpret these imprecise findings as unlikely that there are truly no impacts on time allocation; rather, we believe that we are underpowered to detect changes in allocation over a short 24-hour window.

E Appendix E: Heterogeneous Treatment Effects

In addition to estimating the intent-to-treat effects of our intervention on our entire sample of women, we note that the effects of the family planning intervention may differ for different subgroups. We begin by examining heterogeneous treatment effects of our intervention by women's pregnancy status at baseline, noting that the intervention may have different effects for women who were initially pregnant, and therefore had little need for family planning services at baseline, relative to women who were immediate postpartum at baseline. For the latter group, counseling and the opportunity to visit a clinic using private transport may likely be more salient, given their recent birth and demand to space or limit their next pregnancy. In previous work, we have shown that many of the effects of our intervention have been larger among postpartum women (Karra et al., 2022), and it is reasonable to expect that the effects on labor market outcomes would be similar.

Figure E1 presents findings from the heterogeneity analysis. Specifically, we present the coefficients and confidence intervals of the adjusted ITT effects on labor market outcomes for women and their husbands, stratified by women's pregnancy status at baseline, women's baseline educational attainment, and work status at baseline. The most striking result from this analysis is the differences in the estimated impact of the intervention on labor market outcomes between those women who were pregnant at baseline and those women who were postpartum at baseline. The imprecision in our estimates prevents us from definitively stating that the effects are different between these subgroups; however, the point estimates that we observe suggest that the impact of the intervention on women's labor market outcomes is significantly larger for women who were pregnant at baseline,suggesting that improved access to family planning services allowed women to more effectively re-enter the labor market following their pregnancy.

The differences we observe across the other subgroups are less pronounced. Across education, partitioning our sample seems to reduce our power to detect effects; however, we do detect a large increase in the uptake of wage-earning labor among women who had completed at least a secondary education at baseline, suggesting that women with greater initial human capital attainment were better positioned to benefit from the labor market gains from the intervention. Finally, we see that the effects among women who were not working at baseline largely correlate with the average treatment effects across the entire sample. This may be expected as most of the sample was not working at baseline. However, it also may reflect the persistence of labor market activities throughout and following pregnancy, as women who are already working were unlikely to change their behavior due to the intervention.

F Appendix F: Secondary Work Results

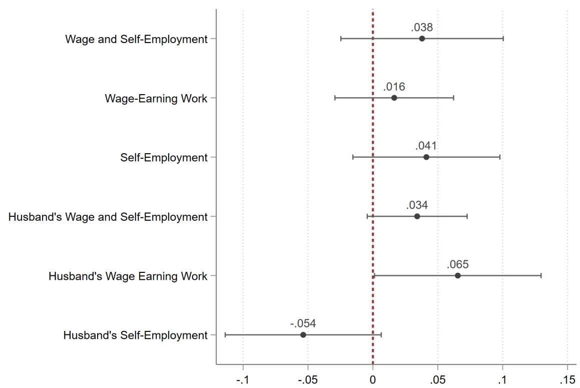
F.1 Ancillary Work Outcomes

In addition to our main outcomes, our survey instruments included questions used to measure agricultural labor, unemployment, and labor force participation. We exclude agricultural labor as an outcome because it is likely not indicative of well-being at the individual level. We exclude unemployment and labor force participation from our main outcomes because the questions used to measure these outcomes were not included in a supplementary phone survey used to survey participants who could not be found in person; data on these outcomes

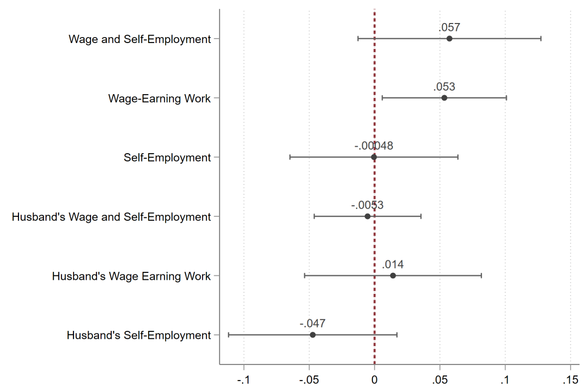
Figure E1: Coefficient Plots of Heterogeneous Effects at Study Endline

(a) Postpartum at Baseline

(b) Pregnant at Baseline



(c) No Secondary Education



(d) Secondary Education

(e) Not Working at Baseline

(f) Working at Baseline

Notes: Points and bracketed lines represent the point estimates and confidence intervals of ITT treatment specifications among the relevant subgroup.

is therefore incomplete.

Nevertheless, in Table F1, we present the effects on these outcome for women in our

Table F1: ITT Effects on Women's Labor Market Outcomes

	(1) Agricultural Work	(2) Unemployment	(3) Labor Force Participation
Unadjusted Estimate	-0.0000360	0.00861	0.0378
	(0.00463)	(0.0113)	(0.0254)
	[1670]	[1515]	[1515]
ANCOVA Estimate	-0.0000336	0.00833	0.0405
	(0.00465)	(0.0114)	(0.0249)
	[1665]	[1507]	[1507]
Fully Adjusted Estimate	-0.0000793	0.00856	0.0369
	(0.00472)	(0.0114)	(0.0249)
	[1663]	[1505]	[1505]

Notes: For all columns, the unit of observation is a woman. The results presented are from OLS models with standard errors in parenthesis and sample sizes in brackets. Results from three estimations are reported; a naïve estimation, an ANCOVA estimation controlling for the baseline value of the outcome variable and a fully adjusted model. Baseline covariates included in the fully adjusted model include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

study sample. Across the outcomes, we find no evidence of program effects. For agricultural work and unemployment, our estimates are precisely measured and very close to zero across our estimations, allowing us to rule out the possibility of economically meaningful program effects. However, although our point estimates on labor force participation are positive and meaningful, our estimates are too imprecise to make declarative statements of program effectiveness. With that said, given the program effects that we have demonstrated on work in the main text, and the fact that we can rule out compensating decreases in unemployment due to the precision of our estimates in column (1), it is reasonable to assume that we are underpowered to detect a true underlying increase in women's labor force participation.

We complement the findings from Table F1 with findings for men in Table F2. Here, we find similarly small and precise estimates on unemployment and agricultural labor, implying that the intervention did not have large effects on these outcomes for men. However, unlike our findings from Table F1, our findings for labor force participation are also close to zero and are more precise than those findings for women. Given that we observe offsetting effects on wage-earning labor and entrepreneurship for men, it may be reasonable to assume, in line with our estimates here, that any potential change in labor force participation among men caused by the intervention is small.

Table F2: ITT Effects on Men's Labor Market Outcomes

	(1) Agricultural Work	(2) Unemployment	(3) Labor Force Participation
Unadjusted Estimate	0.00483	-0.000482	0.0112
	(0.00389)	(0.0103)	(0.00702)
	[1392]	[1392]	[1392]
ANCOVA Estimate	0.00470	-0.000594	0.00962
	(0.00387)	(0.0103)	(0.00681)
	[1384]	[1384]	[1384]
Fully Adjusted Estimate	0.00450	0.000565	0.00931
	(0.00407)	(0.0104)	(0.00676)
	[1383]	[1383]	[1383]

Notes: For all columns, the unit of observation is a man. The results presented are from OLS models with standard errors in parenthesis and sample sizes in brackets. Results from three estimations are reported; a naïve estimation, an ANCOVA estimation controlling for the baseline value of the outcome variable and a fully adjusted model. Baseline covariates included in the fully adjusted model include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

F.2 First-Year Outcomes

In addition to the outcomes we have presented thus far, our study also included a first-year follow-up survey during which our main outcomes were collected. Although we do not present results for this survey in the main text, we present them below in Table F3 for women and Table F4 for men. As one may expect from studying the trends in Figure 1 of the main text, we observe very little by way of differences in outcomes among women during the first-year follow-up. While there is some evidence of improvement in labor outcomes among men when studying our unadjusted estimates, these estimates retain their sign but lose statistical significance when controlling for baseline values of each outcome.

Table F3: ITT Effects on First-Year Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							LN(Labor Income)	
	Wage-Earning or Self-Employment	Wage Earning Work	Self-Employment	Agricultural Work	Unemployment	Labor Force Participation	Extensive Margin	Intensive Margin
Panel A: First-Year Follow up								
Unadjusted Estimate	0.00476	-0.000398	0.00829	0.00272	-0.00557	-0.00245	0.0141	-0.479
	(0.0227)	(0.0148)	(0.0209)	(0.00324)	(0.0106)	(0.0246)	(0.0146)	(0.339)
	[1771]	[1771]	[1771]	[1771]	[1656]	[1656]	[1656]	[159]
ANCOVA Estimate	0.00329	-0.00214	0.0110	0.00244	-0.00470	-0.00323	0.0158	
	(0.0219)	(0.0147)	(0.0200)	(0.00319)	(0.0105)	(0.0236)	(0.0144)	
	[1768]	[1768]	[1768]	[1767]	[1650]	[1650]	[1650]	
Fully Adjusted Estimate	0.000946	-0.00264	0.00878	0.00238	-0.00489	-0.00589	0.0149	-0.475
	(0.0217)	(0.0147)	(0.0199)	(0.00325)	(0.0105)	(0.0234)	(0.0144)	(0.372)
	[1766]	[1766]	[1766]	[1765]	[1648]	[1648]	[1648]	[159]

Notes: For all columns, the unit of observation is a woman. The results presented are from OLS models with standard errors in parenthesis and sample sizes in brackets. Results from three estimations are reported; a naïve estimation, an ANCOVA estimation controlling for the baseline value of the outcome variable and a fully adjusted model. Baseline covariates included in the fully adjusted model include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

Table F4: ITT Effects on First-Year Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							LN(Labor Income)	
	Wage-Earning or Self-Employment	Wage Earning Work	Self-Employment	Agricultural Work	Unemployment	Labor Force Participation	Extensive Margin	Intensive Margin
Panel A: First-Year Follow up								
Unadjusted Estimate	0.0236	0.0422	-0.0163	0.00573	-0.0122	0.0112	0.0480	-0.0665
	(0.0138)	(0.0244)	(0.0241)	(0.00446)	(0.00944)	(0.00658)	(0.0252)	(0.224)
	[1574]	[1574]	[1574]	[1574]	[1574]	[1574]	[1574]	[756]
ANCOVA Estimate	0.0218	0.0265	-0.0151	0.00515	-0.0125	0.0104	0.0297	
	(0.0137)	(0.0213)	(0.0203)	(0.00430)	(0.00949)	(0.00652)	(0.0238)	
	[1569]	[1569]	[1569]	[1569]	[1569]	[1569]	[1569]	
Fully Adjusted Estimate	0.0198	0.0248	-0.0161	0.00545	-0.0123	0.00953	0.0281	-0.0532
	(0.0137)	(0.0213)	(0.0204)	(0.00429)	(0.00956)	(0.00643)	(0.0239)	(0.225)
	[1567]	[1567]	[1567]	[1567]	[1567]	[1567]	[1567]	[756]

Notes: For all columns, the unit of observation is a man. The results presented are from OLS models with standard errors in parenthesis and sample sizes in brackets. Results from three estimations are reported; a naïve estimation, an ANCOVA estimation controlling for the baseline value of the outcome variable and a fully adjusted model. Baseline covariates included in the fully adjusted model include women's age, age of first cohabitation, number of surviving children, ever use of contraception, current use of contraception, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.