Family Planning, Women's Employment, and Household Labor Supply*

David Canning† Mahesh Karra‡ Daniel Maggio§

January 28, 2026

Abstract

We conduct a randomized controlled trial that provides pregnant and postpartum women with comprehensive access to family planning services. We assess the causal impact of the family planning intervention on women's and households' labor market outcomes. We find evidence of a 3.1 to 3.9 percentage point (p.p.) increase in employment and a 2.4 p.p. increase in labor supply among intervention households. These results are primarily driven by a 5.1 p.p. increase in women's employment. An exploration of potential pathways suggests that changes in subsequent fertility, improvements in women's decision-making power over their healthcare, and compensatory shifts in husbands' employment composition together account for roughly half of the variation in the estimated effect on women's employment.

*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.

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

‡Heilbrunn Department of Population and Family Health, Mailman School of Public Health, Columbia University, New York, NY USA. mvk2109@cumc.columbia.edu.

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

1 Introduction

Expanded access to family planning (FP) is widely viewed as a central component of an effective development strategy. From a rights-based perspective, FP promotes women's agency and freedom of choice by strengthening their ability to shape their own reproductive lives (UNFPA, 2014; Newman and Feldman-Jacobs, 2015; Hardee and Jordan, 2021). Beyond its intrinsic value, FP is also understood as a means to many ends, with a substantial body of evidence documenting its significant and positive impacts on fertility and birth spacing, health and human capital, women's work and labor market outcomes, and broader social and economic well-being (Canning and Schultz, 2012; Miller, 2010; Joshi and Schultz, 2007). These links are central to the broader macroeconomic evidence base on the Demographic Dividend, which emphasizes the role of FP in enabling women and couples to effectively realize their fertility preferences by averting unintended pregnancies and mistimed births, thereby facilitating positive fertility change and reducing reproductive uncertainty (Bloom et al., 2003, 2007; Lee and Mason, 2006; Becker and Tomes, 1976; Becker and Lewis, 1973; Cavalcanti et al., 2021). Impacts on fertility and birth spacing, in turn, contribute to investments in women's health and education, as well as in the human capital of their children. Moreover, if labor markets are positioned to absorb additional labor supply, expanded access to FP may also foster female labor force participation and productivity as time and resource constraints are relaxed. Together, these potential impacts suggest that FP may generate economic growth at the aggregate level while simultaneously promoting the welfare of women and their families at the household level (Bloom et al., 2003).

A large quasi-experimental literature has investigated the 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, which is driven in part by access to FP and reproductive health programs, has potentially large and positive effects on women's work. To date, much of the evidence that has directly linked FP access and contraceptive use to women's labor market outcomes has been generated in high-income contexts. Perhaps the most influential works in this literature have investigated the introduction of oral contraceptives ("the pill") on women's labor supply and have shown that the availability of oral contraceptives facilitated the attainment of professional education among women (Goldin and Katz, 2002), lowered women's likelihood of pregnancy early in their lives (Bailey, 2006), and decreased the gender wage gap (Bailey et al., 2012). However, evidence on the link between FP and women's labor market outcomes in low-income

settings is much more limited and mixed. In these settings, the strongest evidence to date has been generated by the Matlab FP, Maternal and Child Health Program (FP-MCH) in Bangladesh, with Schultz (2009) showing that women in program areas earned 40 percent more than women in control areas two decades after the initiation of the program, in part as a result of a 17 percent reduction in fertility. Although this evidence is promising, it warrants three caveats. First, a recent longer-run analysis by Barham et al. (2021) suggests that this effect may have dissipated 35 years after the intervention's initial rollout. Second, the Matlab FP-MCH program combined FP services with other child health services, including a large-scale vaccination campaign, making it difficult to disentangle the labor market effects of FP from other components of the program that were rolled out simultaneously. Finally, we note that the majority of the literature has presumed that the impact of FP on downstream labor market outcomes is mediated exclusively or predominantly through fertility decline. However, in contexts where preferences for birth spacing may outweigh 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 investigate the causal effect of improved access to FP services on women's labor supply and household labor market outcomes using evidence from a randomized controlled trial that was conducted from 2016 to 2019 in urban Malawi. As part of the trial, pregnant or postpartum women were recruited and randomized to a control arm or an intervention arm. Women assigned to the intervention arm received a package of services that included up to six free FP counseling sessions at home, free transportation to a local high-quality FP clinic, and financial reimbursement for all FP commodities and related services at the clinic for a two-year period. Two annual follow-up surveys were conducted following the baseline and rollout of the intervention.

Our study builds on the work of Karra et al. (2022), where we showed that women assigned to the intervention arm were 5.9 percentage points (p.p.) more likely to use contraception and 44 percent less likely to have a subsequent pregnancy within two years of their last birth. In our current analysis, we report intent-to-treat (ITT) estimates on downstream outcomes related to women's employment and household labor supply after two years of exposure to our intervention. We first show that the intervention increased aggregate household employment by 3.1 to 3.9 p.p., which is matched by a 2.4 p.p. increase in household labor supply. These increases are primarily observed in women in the household, who were the target recipients of the FP intervention; these women were 5.1 p.p. more likely to be employed after two years of intervention exposure. Relatedly, we also observe an increase in wage-earning work, where women in the intervention arm are 3.4 p.p. more likely to

report working for a wage compared to women in the control arm. At the same time, we do not observe evidence of intra-household substitution in labor supply between men and women. We also do not observe evidence of an intervention effect on men's employment, unemployment, or labor force participation throughout the study. We do, however, show that husbands of women who were assigned to the intervention arm are 5.4 p.p. less likely to report being self-employed at endline and are 6.3 p.p. more likely to report working for a wage, providing some evidence that they substitute between types of labor.

Our analysis of potential mechanisms centers on the observed gains in women's employment. An examination of plausible pathways suggests that short-term reductions in subsequent fertility, improvements in women's decision-making power over their health care, and compositional changes in partners' employment together account for a substantial share of the variation in the estimated treatment effect. Over the study period, women in the control arm experienced a 15 p.p. increase in employment, likely reflecting re-entry into the labor force following their pregnancy. In contrast, employment among treated women increased by 20 p.p., consistent with the interpretation that delaying or preventing subsequent pregnancies, and thereby reducing uncertainty over future fertility, facilitated this re-entry.

Our paper contributes to two related literatures. First, we provide new experimental evidence on the economic effects of expanded access to FP in a low-income setting. While there is extensive literature showing the positive effects of contraceptive access on women's labor supply in the United States (Goldin and Katz, 2002; Bailey, 2006; Bailey et al., 2012) and other high income settings, causal evidence from low-income contexts remains limited. Literature in this space has generally documented positive effects of FP interventions on first-stage outcomes such as contraceptive use, pregnancy, and birth spacing (Miller and Babiarz, 2016; Mwaikambo et al., 2011).[1] 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 Joshi and Schultz (2007), Miller (2010), and Branson and Byker (2018) have shown that expanded access to FP has positive impacts 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 long-term improvements in their economic well-being. However, the authors noted

[1] A wide range of interventions to improve access to FP - including subsidies that reduce the cost of contraceptive commodities and services (e.g., vouchers or reimbursements), as well as efforts to strengthen supply chains, expand commodity availability at health facilities, and support community-based service delivery - have been shown to increase contraceptive use, reduce pregnancy, and promote birth spacing in settings such as Cameroon (Athey et al., 2023), Malawi (Karra et al., 2022), Bangladesh (Bhatia et al., 1980; Schultz, 2009; Joshi and Schultz, 2007; Debpuur et al., 2002), and rural India (Anukriti et al., 2026). However, more recent evidence from a large-scale experiment in Burkina Faso finds negligible effects of contraceptive subsidies on FP use and fertility outcomes (Dupas et al., 2025).

the challenge of 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 FP-focused intervention that isolates the effects of FP services from complementary interventions that are often bundled with them.

Second, we contribute to the empirical literature on the microfoundations of the Demographic Dividend, which provides a theoretical framework linking FP to fertility and population change, human capital, labor market outcomes, and economic development (Bloom et al., 2003, 2009, 2007; Lee and Mason, 2006; Coale, 1986; Schultz, 1994). Most of the foundational theory and supporting evidence for the Demographic Dividend originate from macroeconomic analyses at the national level, where fertility transitions and shifting population age structures are associated with changes in per capita income, health and human capital, and productivity (Lee and Mason, 2006; Kelley and Schmidt, 1995; Ashraf et al., 2013; Karra et al., 2017). Investigations of the Demographic Dividend in Sub-Saharan Africa have noted the slower pace of fertility decline in the region, which in turn could limit the extent to which a dividend is realized (Bongaarts, 2008; Canning et al., 2015; Caldwell et al., 1992; Casterline and Agyei-Mensah, 2017). Although fertility in Sub-Saharan Africa has fallen more slowly (or, in some countries, plateaued) than in other regions, these aggregate trends obscure significant subnational variation.[2] Exploring these patterns reveals evidence of demographic transitions at these micro levels, suggesting that while aggregate transitions may be slow or uneven, individual and household-level dynamics may still generate economic returns that are aligned with the predicted mechanisms of the Demographic Dividend (Bloom et al., 2012). To date, empirical evidence for the microfoundations of the Demographic Dividend has been primarily drawn from observational or quasi-experimental studies, which estimate the extent to which reductions in fertility and access to contraceptives contribute to women's labor supply (refer to Angrist and Evans (1996), Goldin and Katz (2002), Bailey (2006) as examples). We add to this evidence base by showing that the results we observe may partly be driven by reductions in short inter-pregnancy interval births and greater decision-making power among women who were pregnant at baseline.

The rest of this study is structured as follows. Section 2 provides the conceptual framework that links FP to women's labor market outcomes. Section 3 describes the study setting, 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

[2]For example, although Ethiopia's national total fertility rate remains above 4 children per woman, the capital of Addis Ababa has sustained below-replacement fertility levels for over two decades (Sibanda et al., 2003; Jung, 2023).

mechanisms. We discuss the implications of our findings and conclude in Section 6. We present several secondary analyses in the Appendix.

2 Conceptual 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 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 over future pregnancies and births (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 in particular, 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, 2023). 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, lead to subsequent 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 may not have reached the end of their reproductive years by endline, we cannot directly estimate this relationship in this study. However, we show correlational evidence that short inter-pregnancy intervals significantly reduce the likelihood of women working, and these types of pregnancies decrease as a result of our intervention.

3 Study Design

3.1 Setting: Lilongwe

Our study leverages a randomized controlled trial (RCT) that was conducted in Lilongwe, Malawi, from November 2016 to February 2019.[3] Malawi is one of the most rural countries in the world; as of 2019, 76 percent of Malawi's labor force was employed in agriculture (Our World in Data based on International Labor Organization, 2019). As such, it is worth noting that Lilongwe, the country's capital and urban center, has a labor force that is not representative of the country as a whole. To show this, we leverage data from the 2016 Living Standards and Measurement Survey (LSMS) to describe the compositional differences in women's employment between Lilongwe and the rest of the country (National Statistical Office, 2017).[4] Women's employment rates are substantially lower in Lilongwe than nationally: 41 percent of women in Lilongwe are gainfully employed, compared with 60 percent nationwide. This difference reflects a local labor market that is less concentrated in agriculture. Among employed women in Lilongwe, only 26 percent work on a household farm, compared with 75 percent nationally. Instead, women in Lilongwe are considerably more likely to engage in wage employment (23 percent versus 5 percent nationwide) or to operate a household business (39 percent versus 9 percent nationwide).

In our study, we only ask women about broad categories of labor (e.g., wage work, entrepreneurship). Consequently, when estimating treatment effects on employment, our ability to characterize changes in employment composition is limited. Alternatively, we use the LSMS data to infer the types of jobs women typically hold in this setting. Among wage-employed women in Lilongwe, the most common occupations include hairdressing, sales, teaching, building caretaking, and work in medical or dental services. Together, these five categories make up 63.3 percent of wage employment in Lilongwe. We see that self-employment is much more concentrated; among women who report working for a household business, 76.8 percent report working in some form of retail trade.

These descriptive statistics characterize the types of jobs that women take up when they re-enter the labor market following pregnancy in our study setting. They also underscore the limitations to the external validity of our estimates: our results should likely not be interpreted as representative of the effects of the intervention if it were implemented nationwide. Instead, a more appropriate interpretation may be to consider our findings as more

[3] We provide an abbreviated description of the trial in relation to our current analysis; a more detailed description of the study protocol and intervention implementation is presented elsewhere Karra and Canning (2020).

[4] A table presenting the descriptive statistics described here can be found in Appendix A

externally valid for other urban African contexts, or for settings with labor market structures that are comparable to Lilongwe.

3.2 Study Sample

As part of the trial, 2,143 women[5] 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 package of FP services that consisted of six private, home-based FP counseling sessions, free transportation to a high-quality FP clinic in Lilongwe, and financial reimbursement for all contraceptive methods and related FP services over a two-year period.[6] In Table 1, we present a balance table of labor market characteristics at baseline.We supplement this table with a balance table of the covariates we use in our analysis in Appendix B.[7] 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.[8] 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 may partly signal gender disparities in the labor market, but also reflects the fact that half of the sample women are pregnant and may not be working at baseline.

The data used in this study come from three in-person surveys conducted annually throughout the trial. A baseline survey was conducted from September 2016 to January 2017, and two follow-up surveys were conducted annually with women. Data collection for the first follow-up survey (midline) began in August 2017 and was completed in February 2018. 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.[9] To minimize attrition, abbreviated phone

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

[6]Additional details describing the study sample and protocol can be found in Karra and Canning (2020).

[7]The full balance tables for the trial are presented in Karra et al. (2022) and Maggio et al. (2024).

[8]Note that this differs substantially from the descriptives of women's labor presented in section 3.1 because the women in our sample are either pregnant or less than six months postpartum

[9]In Appendix C, we present an additional attrition balance table for the main outcomes that we discuss in this paper. We see that while the original random assignment is maintained in our endline sample, those lost to follow-up are systematically less likely to be employed across both intervention and control arms. We discuss how we address this threat to external validity in Section 4.2.

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 presented in this study were collected as part of the phone survey; as a result, the analytic samples in this study are smaller than those presented in Karra et al. (2022).[10]

3.3 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 transportation 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 two years. 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.[11] Transportation was provided by a driver who was hired exclusively for the intervention. In addition to the driver, women who used the transportation service were accompanied by a female field manager in the taxi, which served to mitigate any social stigma surrounding a woman traveling alone with another man.

[10] 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 employment, but we cannot infer women's labor force participation, unemployment, or labor market outcomes for their husbands.

[11] 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.

Finally, women assigned to the intervention arm received up to 17,500 MWK ($\sim$ \$25.00 USD) in financial coverage for any costs incurred while receiving FP care at the Kauma Clinic or any other clinic in Lilongwe that they chose to attend. Costs eligible for reimbursement included those related to the procurement and removal of contraceptive methods, FP consultations, treatment and management of contraceptive-related side effects and contraindications, laboratory tests and fees, and other exam fees related to FP and reproductive health care. Although the reimbursement allowance amount was non-transferable for non-FP-related services, women were allowed to redeem any reimbursement amount over multiple visits during 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 effect of our intervention on outcomes related to employment and labor force participation after two years of exposure to our intervention.[12] We present these results at the household level as well as separately for women and men. For our ITT estimates, we run the following standard specification:

$$Y_{mit} = \alpha + \beta_T T_i + \gamma Y_{mi0} + \varepsilon_{mit} \tag{1}$$

where Y_{mit} is the value of the dependent variable for member m of household i at time t. In this study, we use several dependent variables related to the extensive and intensive margins of labor market outcomes and time use, which we discuss in more detail below. In our regressions, T_i represents the treatment assignment for the women living in household i, Y_{mi0} represents the baseline value of the outcome, and ε_{mit} represents the heteroskedastic-robust error term. In our aggregate household models, ε_{mit} is clustered at the household level to account for within-family correlations and dependencies in labor market decisions.

We also estimate the following adjusted ITT specifications:

$$Y_{mit} = \alpha + \beta_T T_i + \gamma Y_{mi0} + \mathbf{X}_{mi0}\zeta + \mathbf{Z}_{mi0}\eta + \varepsilon_{mit} \tag{2}$$

[12] In Appendix D, we report the treatment effect after one year of intervention exposure.

where $\mathbf{X}_{mi0}$ represents the vector of covariates that were used to balance assignment to treatment at baseline, and $\mathbf{Z}_{mi0}$ represents the vector of other baseline covariates that are adjusted for in our regressions.

In our models, the coefficient β_T identifies the ITT effect of being assigned to the intervention arm on outcomes. We choose to study the effect of intervention assignment rather than intervention 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 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." Secondly, studying the ITT 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.[13]

4.2 Sample Attrition

As we note in Section 3, some women in our sample were interviewed by phone, 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, we were able to re-interview 1,672 of our 2,143 sample women, yielding a 78 percent follow-up rate. In contrast, we obtained employment information for only 1,392 husbands, a 65 percent follow-up rate.[14]

Previous analyses have shown that the attrition that we observe seems non-differential between treatment arms in both levels and characteristics (Karra et al., 2022; Maggio et al., 2024). In Appendix C, we supplement these prior analyses by presenting an attrition balance

[13]Nevertheless, it is of great policy interest as to which component of our intervention is likely driving our results. In Maggio et al. (2024), we conducted a LATE analysis using intervention assignment as an instrument for each of the intervention components. While this analysis cannot be interpreted causally because the instrument is under-identified, it does suggest that the per-trip effects of the transportation service are larger than the per-session effects from counseling. However, we observe that uptake of transportation services is substantially lower. Only 211 women out of the 1,043 offered services ever used the transportation component (an average of 2.26 times), while 999 received counseling at some point over the course of the intervention. The gradient in take-up implies that the overall effects of the components are likely similar and highlights a potential policy trade-off between light-touch interventions that can be rolled out more extensively and more intensive, but targeted interventions.

[14]In Maggio et al. (2024), we show that there is a 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.

table of our main outcomes, disaggregated by gender. We first see that, as in prior analyses, the trial's randomization appears robust to the attrition we observe, as households in our endline sample remain balanced in baseline work status. However, we see that women who are lost to follow-up are systematically less likely to be employed and in the labor force relative to women who were recontacted at endline, independent of intervention arm.

Although 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 and selection into our endline survey suggest that it would be important to explore how this attrition might affect our estimates. We investigate the impact of attrition in two ways.

4.2.1 Inverse Propensity Score Weighting

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 sample attrition imbalances that may be correlated with women's education, religion, and employment. 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 treatment effects. We conduct this bounding exercise in two ways. First, in the Kling et al. (2007) tradition, we assume that the missing outcome data are a given number of standard deviations away from the within-intervention arm mean. For these bounds, we choose a relatively conservative 0.1 standard deviation difference. At the lower(upper) end, these bounds explore how we might expect our treatment effects to change if those respondents who attrited in the intervention arm were the least(most) likely to be employed, while those lost to follow-up in the control arm were most(least) likely to be employed.

Second, we present Lee (2009) bounds, which trim the sample proportional to the differential attrition rate.[15] Here, the lower bounds represent how we would expect our treatment effect estimates to evolve if the additional women we were able to survey in the control arm (proportionally) were the least likely to be employed. While the Kling et al. (2007) bounds are agnostic to asymmetry in attrition in either levels or characteristics, Lee (2009) bounds allow us to estimate the treatment effect of the intervention among the subpopulation of

[15] We see from Table C1 that the attrition rates are slightly different across intervention arms, with 20 percent of women lost to follow-up in the control arm, relative to 23.7 percent in the intervention arm. Attrition is more severe but less differential for men, with 34.4 percent lost to follow-up in the control arm, relative to 35.2 percent in the intervention arm.

always-responders. Due to the practical difficulty of including covariates in the Lee (2009) bound methodology, we conduct both of our bounding methods on our unadjusted treatment effects.

5 Results

5.1 ITT Effects

We begin by presenting the effects of our intervention at the individual level, pooling the sample of men and women at endline together. These results are presented in Table 2. Here, we present the unadjusted (Columns 1 & 2) and adjusted (Column 3) ITT estimates for the intervention effect on labor market outcomes at the two-year endline. The ITT effects are presented alongside estimates from the inverse probability weighting model (Column 4), Kling et al. (2007)-type bounds (Columns 5 & 6) and Lee (2009) bounds (Columns 7 & 8).

We observe that members of households in which the woman was assigned to the intervention arm were between 3.1 and 3.9 p.p. (5 and 6 percent) more likely to be employed. In addition, we observe increased participation in wage-earning labor as a result of our intervention, whereby members of these households were between 3.4 and 5.1 p.p. (10 and 16 percent) more likely to report being engaged in wage-earning labor at endline. Both of these estimates are fairly robust to the attrition that we observe. Specifically, they remain largely unchanged when adjusted for propensity score weighting, and in both of our bounding exercises, the results for wage-earning work remain statistically significant and positive. For total employment, our lower bounds lose statistical significance, but remain similar in interpretation. Meanwhile, we see no evidence of a change in self-employment or unemployment as a result of intervention exposure. Finally, we find that the overall labor supply of the household increased by roughly 2.4 p.p. (3.2 percent) over the two-year intervention period. Here, we see that while propensity score weighting increases the precision of our estimate, it may be less robust after accounting for the attrition that we observe, as evidenced by the change of sign in Column (5). Nevertheless, our results suggest that any increases in women's employment and labor supply are not offset by compensating decreases in men's labor market participation.

We decompose our pooled sample results by gender in Table 3, where we first present the results for women only. We find that household-level employment gains are largely driven by women. Specifically, women assigned to the intervention arm are 5.1 p.p. (13.3 percent) more likely to be employed at endline compared to women assigned to the control arm. Similarly, we observe a 3.2 to 3.5 p.p. (27 to 30 percent) increase in wage-earning

employment among women in the intervention arm relative to the control arm.[16] Although our lower bounds are not statistically significant for both outcomes, we note that the range of our bounds is smaller than what we observe in our estimates for the pooled household sample. Similarly to our pooled analysis, we do not observe any evidence of an effect of the intervention on women's self-enterprise or unemployment. Unlike our pooled analysis, our estimates for women's labor force participation are imprecisely measured, although our point estimates suggest between a 2 to 3 p.p. (roughly a 10 percent) increase in women's self-employment and labor force participation as a result of the intervention.

Our pooled and women-stratified findings are promising for understanding the broader labor market implications of our FP intervention. However, we may also expect unintended negative consequences if households view female and male labor supply as substitutes; if so, we may observe a commensurate reduction in male labor market outcomes toward female labor market outcomes as women delay pregnancy and resume working. To explore this possibility, we estimate the treatment effects of our intervention on male employment and labor supply in Table 4. Overall, we do not observe a compensatory decrease in men's labor outcomes, supporting the overall increase in household labor supply that we observe in Table 2. We do, however, find a 5 to 5.4 p.p. (13 to 14 percent) decrease in men's likelihood of self-employment between treatment arms. This decrease, however, seems largely offset by a compensating increase in men's wage-earning labor, although this estimate shrinks when controlling for baseline differences in the outcome. Taken together, these results indicate that as FP services enable women to increase their labor market participation, their husbands spend less time on household enterprises, opting instead for more formal, waged work.

Given the higher attrition that we observe for men, the estimates for men in Table 4 are more sensitive to our bounding analysis. However, in the cases where we observe statistically significant results in men's self-employment and wage-earning work, we see that the bounds that work in the opposite direction of our treatment effects (lower bounds for wage-work and upper bounds for self-enterprise) remain economically meaningful, even if they lose statistical significance.

In Appendix F, we present trends in our main outcomes for the combined sample, men, and women. There, it becomes clear that in our control arm, employment among women increases substantially throughout the study period. However, this increase appears to have accelerated at a faster rate among women in the intervention arm. Given that our sample consists of pregnant and postpartum women at baseline, this likely signals that the interven-

[16] In Appendix E, we aim to substantiate these effects by presenting results for women's time use. Our estimates are noisy, making it difficult to interpret the results, but our point estimates from this analysis suggest that women spend more time on work and less time on childcare and housework.

tion had a role at facilitating women's re-entry into the labor market following pregnancy.

5.2 Potential Mechanisms

In Section 2, we described how we expect FP to affect labor force outcomes through changes in fertility, birth spacing, and control over birth timing. To explore these potential mechanisms, one would typically look to instrument 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. Although we show in Karra et al. (2022) that our intervention leads to a 4 p.p. decrease in short-term fertility, its explanatory power over fertility is low; leveraging our random treatment assignment would therefore yield a weak instrument. In considering a causal mediation analysis, we are unlikely to find a covariate set that satisfies the sequential ignorability assumption, as we expect work and fertility decisions to be highly correlated and therefore confounded (Imai et al., 2010).

In the absence of a viable causal strategy to identify the mediating pathways underlying our results, in Table 5 we explore the extent to which controlling for subsequent fertility, empowerment over her healthcare, and husband's employment composition attenuates our observed treatment effect on women's overall employment. To our main treatment effect estimate in equation 1, we sequentially add a dummy variable that takes the value one if a woman had a subsequent pregnancy following her index pregnancy that made her eligible for the trial, and a dummy variable that takes the value one if a woman reports to have control over her healthcare decisions,[17] and two dummy variables that take on values of one if the respondent's husband is self-employed and if the respondent's husband is engaged in waged work. To ensure the samples are comparable across estimations, missing covariates are coded to zero, and a dummy variable that takes the value one if the data were originally missing is included.

In Column (2), we see that subsequent pregnancies are negatively correlated with women's employment; regardless of treatment arm, women who get pregnant again during our two-year observation window are 18.6 percentage points less likely to be employed than their peers who did not have subsequent pregnancies. We also see that, consistent with our theoretical framework, controlling for these subsequent pregnancies attenuates our treatment effect. This suggests that the change in short-term fertility resulting from the intervention may explain at least part of the effect that we observe. However, given that pregnancies are a

[17] We use the following survey question that was asked of women to code this variable: "Who usually makes decisions about health care for yourself: you, your husband/partner, you and your husband/partner, jointly, or someone else?" We code this variable to take the value one if the respondent reports making the decision either alone or jointly with her partner, and zero otherwise.

rare event in our sample, this attenuation still leaves 80 percent of the treatment effect that is not explained by the change in fertility.

To the extent that healthcare decision-making serves as a proxy for women's empowerment, Column (3) suggests that this empowerment is positively associated with employment. Specifically, women who have a say over their healthcare decisions are 16.5 percentage points more likely to be employed than those who do not. Adding this covariate to our ITT estimation attenuates our result by an additional 9 percent. However, it is worth noting that in Column (3), although the estimate is attenuated by 29 percent relative to our main estimate, we can no longer reject the null hypothesis at the 10 percent level.

Finally, given our results in Table 4, we may believe that it is specifically among households where men reallocate their labor that women are able to increase their employment. However, in Column (4), we see no evidence of an association between men's employment status and women's employment. However, controlling for these variables does further attenuates our measured treatment effect, resulting in a coefficient that is only 53 percent of our original treatment effect estimate.

The results presented in Table 5 are consistent with the theoretical framework presented in Section 2 inasmuch as we can explain changes in employment through changes in fertility, empowerment, and intra-household labor reallocation. However, the fact that we are left with a positive (albeit insignificant) treatment effect on employment after including these covariates suggests that there may be scope for additional pathways that we are unable to measure. One potential explanation for this is that the effects of fertility change on women's employment likely operate through multiple channels. For example, when pregnancies are realized, they place constraints on women's time. However, the increased certainty that women have about their fertility horizon as a result of changes in FP behavior is likely to also affect how they engage in the labor market. Realized pregnancies are likely to be an imperfect proxy for this latent certainty due to their limited frequency. As such, it is possible that with a measure that is more reflective of women's priors of their likelihood of getting pregnant, we would be able to explain a larger proportion of the variation in the observed treatment effect.

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 ITT estimates show large labor market gains among households that received the intervention. Members of these households are up to 3.9 p.p. more likely to be employed two years

into the intervention (a 5 percent relative increase). These effects are primarily driven by women, who were 5.1 p.p. (15 percent) 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. Furthermore, while we observe some evidence of compositional changes in men's employment status, we do not observe deleterious effects on their overall employment or labor market participation, 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 have recently come into question. Our results contrast with those of Barham et al. (2021), who found no 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 previous experimental findings (Karra et al., 2022; Maggio et al., 2024), the results we present are part of a more extensive body of evidence that makes the case for improving access to high-quality FP and reproductive health services for 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 underestimate the impact of, and therefore 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. 5

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. 40

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, 4, 16

Anukriti, S., C. Herrera-Almanza, and M. Karra (2026, March). Bring a friend: Leveraging financial and peer support to improve women's reproductive agency in India. *Journal of Development Economics 180*, 103706. 3

Ashraf, Q. H., D. N. Weil, and J. Wilde (2013). The Effect of Fertility Reduction on Economic Growth. *Population and Development Review 39*(1), 97–130. Publisher: [Population Council, Wiley]. 4

Athey, S., K. Bergstrom, V. Hadad, J. C. Jamison, B. Özler, L. Parisotto, and J. D. Sama (2023, October). Can personalized digital counseling improve consumer search for modern contraceptive methods? *Science Advances 9*(40), eadg4420. Publisher: American Association for the Advancement of Science. 3

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, 4

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, 16

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), e2101160118. 2, 3, 16

Becker, G. (1960). An Economic Analysis of Fertility. NBER Chapters, National Bureau of Economic Research, Inc. 5

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. 5

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, 4

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay (2009). Fertility, Female Labor Force Participation, and the Demographic Dividend. *Journal of Economic Growth 14*, 79–101. 4, 5

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay (2012, August). Microeconomic Foundations of the Demographic Dividend. Technical report, Program on the Global Demography of Aging, Cambridge, MA. 4

Bloom, D. E., D. Canning, R. K. Mansfield, and M. Moore (2007). Demographic Change, Social Security Systems and Savings. *Journal of Monetary Economics 54*, 92–114. 4

Bongaarts, J. (2008). Fertility Transitions in Developing Countries: Progress or Stagnation? *Studies in Family Planning 39*(2), 105–110. 4

Branson, N. and T. Byker (2018, January). Causes and consequences of teen childbearing: Evidence from a reproductive health intervention in South Africa. *Journal of Health Economics 57*, 221–235. 3

Caldwell, J. C., I. O. Orubuloye, and P. Caldwell (1992, June). Fertility Decline in Africa: A New Type of Transition? *Population and Development Review 18*, 211–242. 4

Canning, D., S. Raja, and A. S. Yazbeck (Eds.) (2015, October). *Africa's Demographic Transition: Dividend or Disaster?* The World Bank. 4

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

Casterline, J. B. and S. Agyei-Mensah (2017). Fertility Desires and the Course of Fertility Decline in sub-Saharan Africa. *Population and Development Review 43*, 84–111. Publisher: [Population Council, Wiley]. 4

Cavalcanti, T., G. Kocharkov, and C. Santos (2021). Family Planning and Development: Aggregate Effects of Contraceptive Use. *The Economic Journal 131*(634), 624–657. 1, 5

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, 16

Coale, A. J. (1986). *The decline of fertility in Europe since the eighteenth century as a chapter in demographic history*. Princeton University Press. 4

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, 16

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

Dupas, P., S. Jayachandran, A. Lleras-Muney, and P. Rossi (2025). The Negligible Effect of Free Contraception on Fertility: Experimental Evidence from Burkina Faso. *American Economic Review 115*(8), 2659–2688. 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. 5

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, 4

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. 14

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

Jung, M. (2023, September). *Spatial aspects of fertility change in Ethiopia between 2000 and 2016: a district-level analysis*. doctoral, London School of Hygiene & Tropical Medicine. 4

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. 6, 7

Karra, M. and D. Canning (2023). Hormonal Contraception and Hemoglobin Levels in Urban Malawi: A Longitudinal Analysis. *Working Paper*. 5

Karra, M., D. Canning, and J. Wilde (2017). The Effect of Fertility Decline on Economic Growth in Africa: A Macrosimulation Model. *Population and Development Review 43*, 237–263. Publisher: [Population Council, Wiley]. 4

Karra, M., D. Maggio, M. Guo, B. Ngwira, and D. Canning (2022). 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. 2, 3, 5, 7, 8, 10, 14, 16, 28

Kelley, A. C. and R. M. Schmidt (1995). Aggregate population and economic growth correlations: the role of the components of demographic change. *Demography 32*(4), 543–555. 4

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. 11, 12

Kohler, H.-P. (2012). Copenhagen Consensus 2012: Challenge Paper on "Population Growth". pp. 79. 16

Kohler, H.-P. and M. Karra (2019). The Benefits and Costs of Scaling Up Family Planning in Africa. Technical report, Copenhagen Consensus Center, Copenhagen, Denmark. 16

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies 76*(3), 1071–1102. Publisher: Oxford University Press. 11, 12

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, 4

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. 11

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. 7, 10, 11, 16, 28

Miller, G. (2010). Contraception as Development? New Evidence from Family Planning in Colombia. *The Economic Journal 120*(545), 709–736. Publisher: [Royal Economic Society, Wiley]. 1, 3

Miller, G. and K. S. Babiarz (2016, March). Family Planning Program Effects: Evidence from Microdata. *Population and Development Review 42*(1), 7–26. 3

Mwaikambo, L., I. S. Speizer, A. Schurmann, G. Morgan, and F. Fikree (2011). What works in family planning interventions: A systematic review. *Studies in Family Planning 42*(2), 67–82. 3

National Statistical Office (2017). Fourth Integrated Household Survey 2016-2017. 6

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

Our World in Data based on International Labor Organization (2019). Share of the labor force employed in agriculture. 6

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. 5

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. 5

Schultz, T. P. (1994). Human Capital, Family Planning, and Their Effects on Population Growth. *The American Economic Review 84*(2), 255–260. Publisher: American Economic Association. 4

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

Sibanda, A., Z. Woubalem, D. P. Hogan, and D. P. Lindstrom (2003). The Proximate Determinants of the Decline to Below-replacement Fertility in Addis Ababa, Ethiopia. *Studies in Family Planning 34*(1), 1–7. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1728-4465.2003.00001.x. 4

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.

Table 2: ITT Effects on Household Work, Year 2

					KLK Bounds		Lee Bounds	
	(1) Unadjusted	(2) ANCOVA	(3) Adjusted Estimate	(4) IPW	(5) Lower	(6) Upper	(7) Lower	(8) Upper
Panel A: Employment								
Treatment	0.0394***	0.0330**	0.0310**	0.0346**	0.0122	0.0667***	0.0167	0.0532***
	(0.0149)	(0.0149)	(0.0150)	(0.0150)	(0.0124)	(0.0124)	(0.0216)	(0.0192)
Control Mean	0.623	0.623	0.623					
Panel B: Wage-Earning Employment								
Treatment	0.0516***	0.0359**	0.0339**	0.0369**	0.0245**	0.0787***	0.0398**	0.0763***
	(0.0152)	(0.0143)	(0.0143)	(0.0144)	(0.0123)	(0.0123)	(0.0187)	(0.0222)
Control Mean	0.323	0.323	0.323					
Panel C: Self-Enterprise								
Treatment	-0.0113	-0.00855	-0.00881	-0.00703	-0.0378***	0.0153	-0.0231	0.0134
	(0.0171)	(0.0156)	(0.0156)	(0.0156)	(0.0121)	(0.0121)	(0.0184)	(0.0219)
Control Mean	0.324	0.324	0.324					
Panel D: Unemployment								
Treatment	0.00429	0.00422	0.00487	0.00467	-0.00900*	0.0176***	0.00363	0.0190
	(0.00767)	(0.00770)	(0.00772)	(0.00779)	(0.00523)	(0.00523)	(0.00780)	(0.0219)
Control Mean	0.043	0.043	0.043					
Panel E: Labor Force Participation								
Treatment	0.0241*	0.0257*	0.0242*	0.0278**	-0.00294	0.0511***	0.0124	0.0278*
	(0.0135)	(0.0135)	(0.0135)	(0.0137)	(0.0106)	(0.0106)	(0.0226)	(0.0166)
Control Mean	0.758	0.758	0.758					
Observations	3064	3052	3049	3049	4280	4280	4280	4280

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

Notes: For all columns, the unit of observation is a respondent. The results presented are from OLS models with standard errors in parentheses. OLS estimates are accompanied by estimates from an inverse probability weighting model (Column 4), Kling-Leibman-type bounds (Columns 5 & 6), and Lee-type bounds (Columns 7 & 8). The adjusted regressions (Column 3) 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. Standard errors are presented in parentheses and clustered at the household level.

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

					KKL Bounds		Lee Bounds	
	(1) Unadjusted	(2) ANCOVA	(3) Adjusted Estimate	(4) IPW	(5) Lower	(6) Upper	(7) Lower	(8) Upper
Panel A: Employment								
Treatment	0.0515**	0.0512**	0.0479**	0.0511**	0.0299	0.0731***	0.0323	0.0825***
	(0.0241)	(0.0237)	(0.0237)	(0.0237)	(0.0188)	(0.0188)	(0.0263)	(0.0288)
Control Mean	0.382	0.382	0.382					
Panel B: Wage-Earning Employment								
Treatment	0.0353**	0.0338**	0.0324*	0.0339**	0.0203	0.0503***	0.0295*	0.0796***
	(0.0168)	(0.0166)	(0.0169)	(0.0168)	(0.0131)	(0.0131)	(0.0174)	(0.0274)
Control Mean	0.117	0.117	0.117					
Panel C: Self-Enterprise								
Treatment	0.0238	0.0265	0.0247	0.0276	0.00394	0.0437**	0.0100	0.0602**
	(0.0222)	(0.0216)	(0.0217)	(0.0216)	(0.0173)	(0.0173)	(0.0236)	(0.0287)
Control Mean	0.274	0.274	0.274					
Panel D: Unemployment								
Treatment	0.00861	0.00833	0.00856	0.00869	-0.00425	0.0215***	0.00810	0.0189
	(0.0113)	(0.0114)	(0.0114)	(0.0115)	(0.00803)	(0.00803)	(0.0115)	(0.0291)
Control Mean	0.0468	0.0468	0.0468					
Panel E: Labor Force Participation								
Treatment	0.0378	0.0405	0.0369	0.0426*	0.00892	0.0667***	0.0318	0.0426
	(0.0254)	(0.0249)	(0.0249)	(0.0251)	(0.0180)	(0.0180)	(0.0299)	(0.0284)
Control Mean	0.556	0.556	0.556					
Observations	1672	1668	1666	1666	2140	2140	2140	2140

*** 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 parentheses. OLS estimates are accompanied by estimates from an inverse probability weighting model (Column 4), Kling-Leibman-type bounds (Columns 5 & 6), and Lee-type bounds (Columns 7 & 8). The adjusted regressions (Column 3) 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: ITT Effects on Men's Work, Year 2

					KLK Bounds		Lee Bounds	
	(1) Unadjusted	(2) ANCOVA	(3) Adjusted Estimate	(4) IPW	(5) Lower	(6) Upper	(7) Lower	(8) Upper
Panel A: Employment								
Treatment	0.0169	0.0161	0.0148	0.0150	-0.00151	0.0353***	-0.00186	0.0186
	(0.0141)	(0.0141)	(0.0142)	(0.0140)	(0.00919)	(0.00919)	(0.0329)	(0.0145)
Control Mean	0.917	0.917	0.917					
Panel B: Wage-Earning Employment								
Treatment	0.0637**	0.0430*	0.0397*	0.0434*	0.0296*	0.0979***	0.0520	0.0725**
	(0.0262)	(0.0237)	(0.0237)	(0.0240)	(0.0170)	(0.0170)	(0.0323)	(0.0298)
Control Mean	0.573	0.573	0.573					
Panel C: Self-Enterprise								
Treatment	-0.0545**	-0.0520**	-0.0503**	-0.0516**	-0.0880***	-0.0210	-0.0623**	-0.0419
	(0.0257)	(0.0223)	(0.0224)	(0.0227)	(0.0167)	(0.0167)	(0.0288)	(0.0327)
Control Mean	0.386	0.386	0.386					
Panel D: Unemployment								
Treatment	-0.000482	-0.000564	0.000565	-0.000149	-0.0139**	0.0129*	-0.00127	0.0192
	(0.0103)	(0.0103)	(0.0104)	(0.0105)	(0.00669)	(0.00669)	(0.0104)	(0.0328)
Control Mean	0.038	0.038	0.038					
Panel E: Labor Force Participation								
Treatment	0.0112	0.00962	0.00931	0.00877	0.00208	0.0202***	-0.00882	0.0116
	(0.00702)	(0.00681)	(0.00676)	(0.00673)	(0.00458)	(0.00458)	(0.0324)	(0.00714)
Control Mean	0.977	0.977	0.977					
Observations	1392	1384	1383	1383	2140	2140	2140	2140

*** 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 parentheses. OLS estimates are accompanied by estimates from an inverse probability weighting model (Column 4), Kling-Leibman-type bounds (Columns 5 & 6), and Lee-type bounds (Columns 7 & 8). The adjusted regressions (Column 3) 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 5: Analysis of Potential Mechanisms

	(1)	(2)	(3)	(4)
Treatment	0.0512**	0.0412*	0.0367	0.0273
	(0.0237)	(0.0237)	(0.0234)	(0.0229)
Subsequent Pregnancy		-0.186***	-0.196***	-0.187***
		(0.0406)	(0.0403)	(0.0398)
Involvement in Healthcare Decisions			0.165***	0.169***
			(0.0251)	(0.0246)
Husband's Self Employment				-0.0281
				(0.0408)
Husband's Wage Work				-0.0419
				(0.0401)
Observations	1668	1668	1668	1668

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

Notes: For all columns, the outcome variable is women's employment, and the unit of observation is a woman. Heteroskedasticity-robust standard errors are reported in parentheses. To preserve a comparable sample across estimations, missing covariates are coded to zero, and a dummy variable is included that takes the value one if the covariate was originally missing.

A Appendix A: LSMS Descriptive Statistics

Table A1: Employment Descriptives from Living Standards and Measurement Surveys

	Lilongwe	Malawi
Number of Women		
	1,259 Women	28,086 Women
Of those women, . . .		
	40.7 percent (513) are economically active	59.5 percent (16,687) are economically active
Of the economically active women. . .		
	23 percent (118) work for a wage	4.59 percent (766) work for a wage
	38.6 percent (198) work in a household business	8.97 percent (1,496) work in a household business
	26.12 percent (134) work on a household farm	75.37 percent (12,577) work on a household farm
	2 work as an apprentice	30 work as an apprentice
	11.9 percent (61) rely on Ganyu	10.9 percent (1,818) rely on Ganyu
Among wage earners, the top 5 jobs are. . .		
	Hairdressers (21.9 percent)	Teachers (17.9 percent)
	Saleswomen or shop assistants (15.4 percent)	Maids and Housekeepers (17.6 percent)
	Teachers (12.2 percent)	Medical or dental-related workers (10.8 percent)
	Buildings caretakers and related workers (7.3 percent)	Agricultural and animal husbandry workers (8.9 percent)
	Medical or dental-related workers (6.5 percent)	Saleswomen or shop assistants (8.3 percent)
Of the household enterprises, the top 5 sectors are. . .		
	Retail trade, except of motor vehicles (76.8 percent)	Retail trade, except of motor vehicles (64.1 percent)
	Repair of computers and personal goods (3.4 percent)	Processing and Manufacturing of food products (7.1 percent)
	Land transport and transport via pipelines (3.4 percent)	Forestry and logging (5.8 percent)
	Other personal service activities (eg. Hairdressing) (2.5 percent)	Land transport and transport via pipelines (4.4 percent)
	Distilling, rectifying and blending of spirits (2.1 percent)	Distilling, rectifying and blending of spirits (4.0 percent)

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; as such, the test of balance is presented at the baseline value for the full sample.

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.

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

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

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

C Appendix C: Attrition Balance

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								
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 Employment	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								
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 Employment	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.

D Appendix D: First-Year Work Results

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 D1 for women and Table D2 for men. 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 D1: ITT Effects on First-Year Outcomes, Women

	(1) Employment	(2) Wage Earning Work	(3) Self-Employment	(4) Unemployment	(5) Labor Force Participation
Panel A: First-Year Follow up					
Unadjusted Estimate	0.00476	-0.000398	0.00829	-0.00557	-0.00245
	(0.0227)	(0.0148)	(0.0209)	(0.0106)	(0.0246)
	[1771]	[1771]	[1771]	[1656]	[1656]
ANCOVA Estimate	0.00329	-0.00214	0.0110	-0.00470	-0.00323
	(0.0219)	(0.0147)	(0.0200)	(0.0105)	(0.0236)
	[1768]	[1768]	[1768]	[1650]	[1650]
Fully Adjusted Estimate	0.000946	-0.00264	0.00878	-0.00489	-0.00589
	(0.0217)	(0.0147)	(0.0199)	(0.0105)	(0.0234)
	[1766]	[1766]	[1766]	[1648]	[1648]

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 na "ive 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 D2: ITT Effects on First-Year Outcomes, Men

	(1) Employment	(2) Wage Earning Work	(3) Self-Employment	(4) Unemployment	(5) Labor Force Participation
Panel A: First-Year Follow up					
Unadjusted Estimate	0.0236	0.0422	-0.0163	-0.0122	0.0112
	(0.0138)	(0.0244)	(0.0241)	(0.00944)	(0.00658)
	[1574]	[1574]	[1574]	[1574]	[1574]
ANCOVA Estimate	0.0218	0.0265	-0.0151	-0.0125	0.0104
	(0.0137)	(0.0213)	(0.0203)	(0.00949)	(0.00652)
	[1569]	[1569]	[1569]	[1569]	[1569]
Fully Adjusted Estimate	0.0198	0.0248	-0.0161	-0.0123	0.00953
	(0.0137)	(0.0213)	(0.0204)	(0.00956)	(0.00643)
	[1567]	[1567]	[1567]	[1567]	[1567]

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 na "ive 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.

E Appendix E: Women's Time Use

Table E1: 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 na "ive treatment effect on outcomes, adjusted only by the value of the outcome at baseline. Heteroskedasticity-robust standard errors are reported.

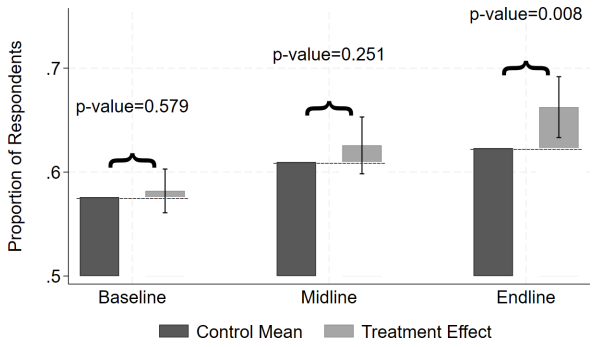
In Section 2 of the main text, we outlined potential reasons why we may expect family planning services to improve women's labor market outcomes, citing the relaxation of constraints from childcare and other related responsibilities that women face. Within our data, we can test this hypothesis by presenting the effects of our intervention on women's time use.[18] In Table E1, we present these results for the first-year follow-up and the second-year endline. 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.

[18]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.

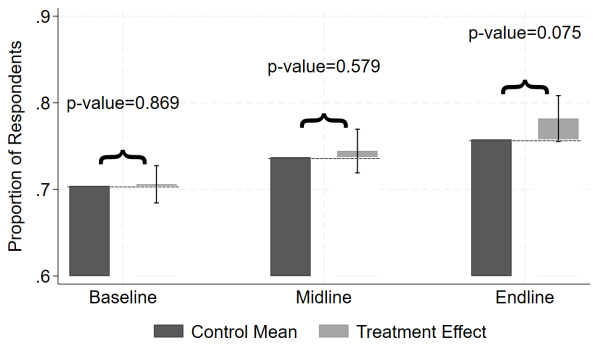
F Appendix F: Trends in Main Outcomes

The main results of our analysis are summarized in Figure F1, which presents intervention impacts on employment and labor force participation over time for households, women, and men. Panels (a) and (b) show that employment and labor force participation are balanced across intervention and control arms at baseline and monotonically increase across survey waves in both arms, with intervention effects emerging at endline.[19] We see similar trends for women, while men's labor market outcomes remain fairly stable. Although men's outcomes appear to improve relative to the control arm, no evidence emerges of a significant improvement in outcomes.

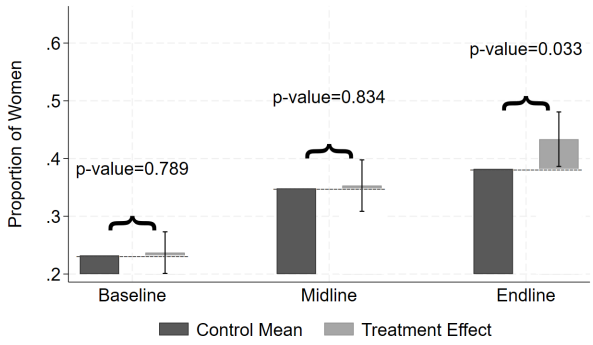
[19]The monotonic improvement in outcomes across groups throughout the intervention period is likely reflective of women non-differentially reentering the labor market following their pregnancy.



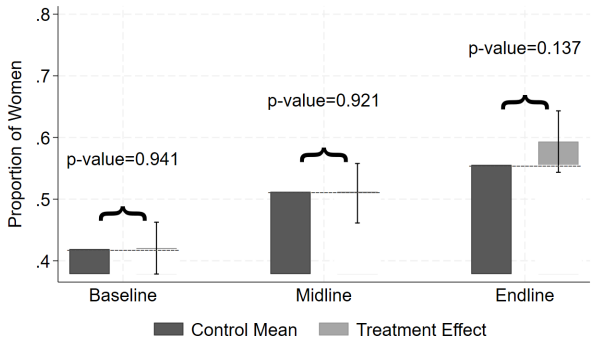
(a) Total Employment



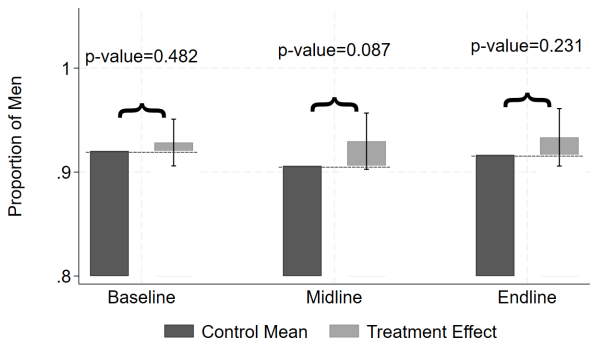
(b) Total Labor Force Participation



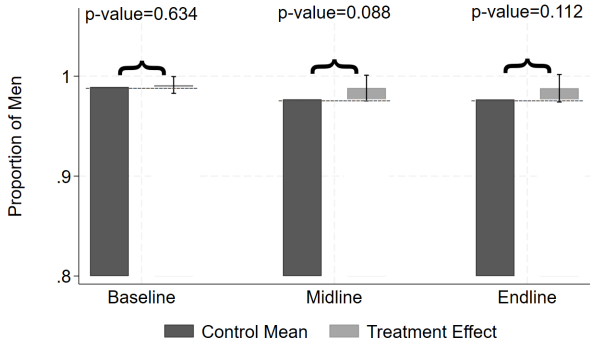
(c) Women's Employment



(d) Women's Labor Force Participation



(e) Men's Employment



(f) Men's Labor Force Participation

Figure F1: Work Outcome Trends over Time

Notes: In each figure, the unit of observation is an individual. In panels (c) and (d), these individuals are restricted to women. In panels (e) and (f), they are restricted to men. The p-values reported in each figure represent the p-value on the test of significance of the unadjusted treatment effect estimate during each survey wave of our study.

G Appendix G: Heterogeneous Effects

We present ITT results beyond the gender-stratified analysis by stratifying our aggregate effects at the household, woman, and man level by women's baseline pregnancy status, age, and educational attainment. Figure G1 presents the differences between groups for each of our main outcomes at the household, woman, and man levels. This figure is accompanied by Table G1, which reports the p-value associated with each test.

We begin by focusing on household-level heterogeneity. Here, we see that none of the group differences are significant at the 5 percent level. However, we observe suggestive evidence that the intervention effects are larger for individuals living in households where the enrolled woman was younger and pregnant at baseline. This is especially true for overall employment and self-employment, where individuals living in households with older (over 25) women at baseline experience a 5 p.p. smaller treatment effect. Meanwhile, individuals living in households with pregnant women and younger women experience a 5 p.p. larger treatment effect.

This pattern holds when looking at the woman only sample. Here, while we are still slightly underpowered to detect group differences in the treatment effect, we see that the observed effects are heavily concentrated among pregnant women. These women experience roughly a 10 p.p. increase in employment, wage-earning work, and labor force participation. Meanwhile, we observe no evidence of a treatment effect among women who were postpartum at baseline.

Finally, the only outcome for which we observe strong evidence of heterogeneity in the treatment effect for men is in self-employment. Specifically, we observe a symmetric shift similar to that observed for women's wage-earning labor in households where women were pregnant at baseline. This would suggest that among these households, there is a compensatory response to women's labor market entry in these households.

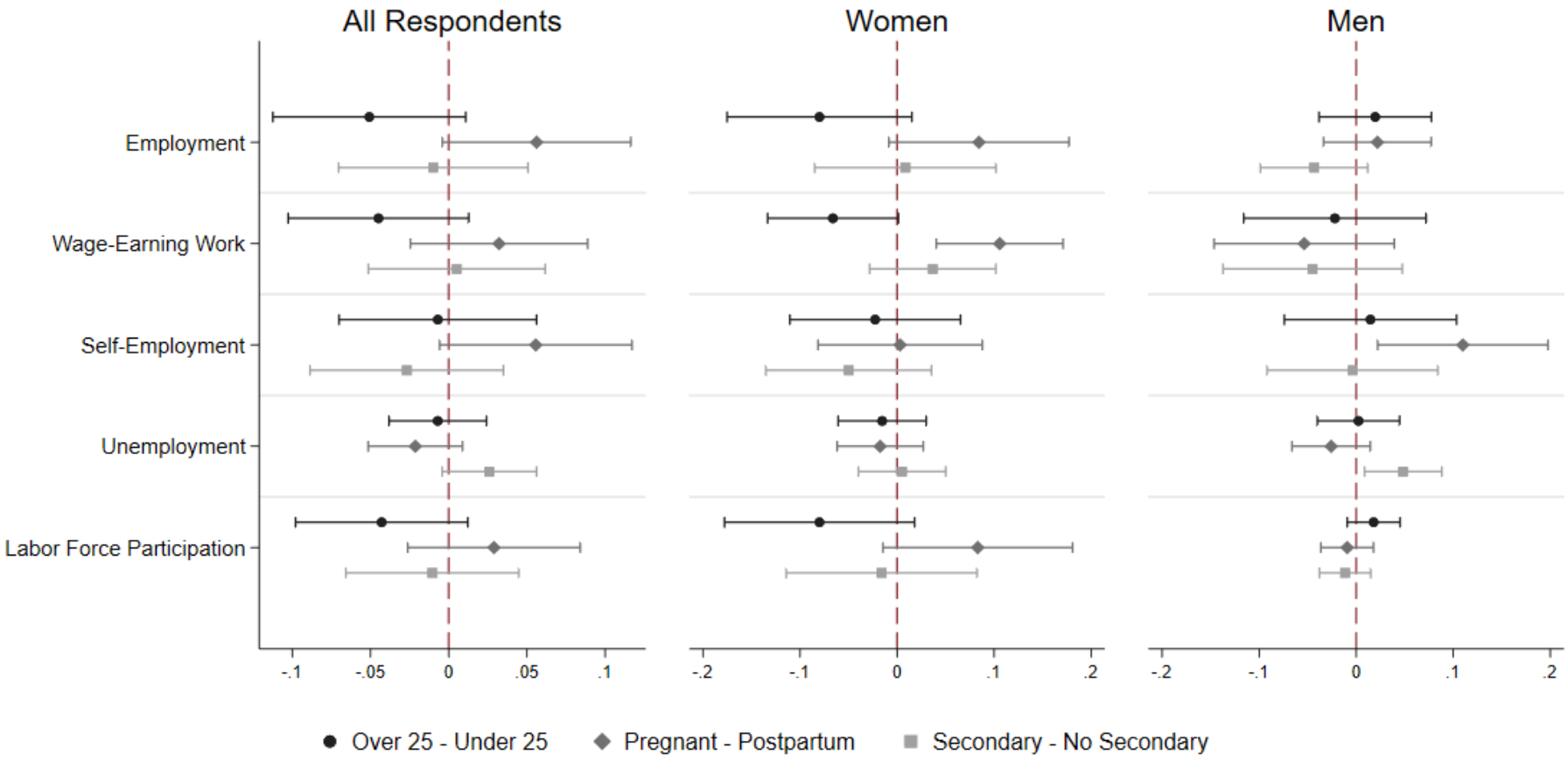


Figure G1: Heterogeneous Effects, Year 2

Notes: In each panel, the unit of observation is an individual. In the middle panel, the sample is restricted to women. In the right panel, it is restricted to men. Each point represents the difference in the measured treatment effect estimates across groups. For example, the "Pregnant - Postpartum" points present the treatment effect among women who were pregnant at baseline, minus those who were postpartum. Standard errors are clustered at the household level for estimations that include both men and women. Estimations that include only men or only women use heteroskedasticity-robust standard errors.

Table G1: P-Values for Tests of Group Differences

	Joint	Women	Men
Employment			
Age	0.107	0.100	0.505
Pregnancy	0.067	0.075	0.438
Education	0.750	0.857	0.125
Wage Earning Employment			
Age	0.127	0.055	0.647
Pregnancy	0.265	0.002	0.258
Education	0.858	0.268	0.341
Self-Employment			
Age	0.829	0.615	0.746
Pregnancy	0.076	0.941	0.014
Education	0.395	0.252	0.934
Unemployment			
Age	0.660	0.507	0.910
Pregnancy	0.165	0.442	0.210
Education	0.092	0.821	0.017
Labor Force Participation			
Age	0.127	0.110	0.194
Pregnancy	0.304	0.095	0.502
Education	0.711	0.752	0.398

Notes: Table reports p-values accompanying the test of differences in coefficients reported in Table G1. Each value accompanies one coefficient reported in that figure and maps directly in terms of location and label.

H Appendix H: 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*.