Family Planning and Women's Labor Supply: Experimental Evidence from Urban Malawi*

Mahesh Karra† Daniel Maggio‡ David Canning§

July 17, 2023

PRELIMINARY - DO NOT CITE

Abstract

We conducted a randomized controlled trial that provided pregnant and immediate postpartum women with improved access to family planning through counseling, free transport to a clinic, and financial reimbursement for family planning services. We assess the causal impact of our intervention on women's labor market outcomes and find that women are 5.2 percentage points more likely to be employed after two years of intervention exposure, which is driven by a 3.38 percentage point increase in women's participation in wage-earning labor. The intervention resulted in a higher proportion of women reporting that they earn labor income, although we do not find evidence of increases in income levels among women who report earnings. Among women's husbands, we find evidence of substitution away from self-enterprise and towards wage-earning labor but no evidence of changes in overall labor force participation, implying that the overall household labor supply increased as a result of our intervention. Our results suggest that the positive effects of improved access to family planning extend beyond fertility and health to women's labor supply.

*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 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.

‡Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY USA. dwm272@cornell.edu.

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

1 Introduction

Ever since Thomas Malthus penned his *Essay on the Principle of Population*, social scientists have been concerned about the economic consequences of high population growth (Malthus et al., 1992). However, as time has proven fears of the *Malthusian catastrophe* to be overstated, much of the focus on fertility has shifted towards the economic *benefits* of decreasing population growth. This "demographic dividend" is particularly salient for countries in the middle stages of the demographic transition. During these middle stages, fertility declines lead to a rapid increase in labor market participation and, in turn, per capita income (Bloom et al., 2003, 2007; Lee and Mason, 2006).

While much of the excitement about the demographic dividend is derived from the potential macroeconomic consequences of demographic transition, these are inherently microeconomic relationships. Namely, as fertility within a household decreases, the household's time constraint is relaxed, allowing its members to increase their labor supply (Becker, 1960). In this sense, the existence of a demographic dividend hinges on the causal relationship between fertility and household labor supply.

Although this relationship may seem mechanical, there are numerous challenges in demonstrating such a causal link. The economic theory of fertility leads us to believe that decreases in fertility *should* cause increases in household labor supply. However, it also tells us that rising incomes increase the opportunity cost of having children, thus reducing fertility and creating a challenge of dual causality in causal estimation. To circumvent this challenge, previous researchers have exploited quasi-random variation in household fertility created by non-singleton births (Rosenzweig and Wolpin, 1980) and by sex-selective childbearing (Angrist and Evans, 1996; Chun and Oh, 2002; Cruces and Galiani, 2007) to explore the effects of fertility on maternal labor supply. This literature is generally consistent in that decreases in fertility have large positive effects on female labor supply. However, there is a relative dearth of experimental evidence on the topic.

In this study, we estimate the causal effect of improved access to family planning (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, 2,143 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. This paper reports intent-to-treat

(ITT) estimates of improved access to FP on women's labor market outcomes after two-years of exposure to our intervention. We show that women who received the comprehensive FP package were 5.2 percentage points (p.p.) more likely to be gainfully employed[1] two years following the intervention. These changes seem to be largely driven by an increase in wage-earning employment, and we show that women in the intervention arm are 3.8 p.p. more likely to report earning labor income compared to women in the control arm. However, conditional on earning an income, we do not find that women earn more due to the intervention. 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 employment, unemployment, or labor force participation throughout the study. We do, however, show that husbands are 5.3 p.p. less likely to report being self-employed at endline, providing some evidence that they substitute self-employment in favor of wage-earning labor. These results on both women's and men's labor market outcomes suggest that the intervention increased the overall labor supply of the household.

Our paper makes two key contributions. Through our experimental design, we are able to more effectively identify the causal links between fertility 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 contribute to women's employment and labor supply (refer to Angrist and Evans (1996) as an example). We add to this evidence base with our experimental findings, where we leverage our family planning intervention as an instrument for fertility. Our second contribution identifies the causal effects of family planning more broadly. Literature in this space has, to date, provided a range of evidence to highlight the positive impacts of family planning 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 that links family planning to more downstream outcomes related to women's social and economic well-being, and evidence of the economic impacts of family planning is mixed. On the one hand, studies by XX and XX have shown positive impacts of family planning on women's labor market outcomes. In contrast, recent work by Barham et al. (2021) found that women exposed to the longstanding Matlab Maternal and Child Health and Family Planning program in Bangladesh experienced no significant improvements in their economic well-being, although the authors noted the challenge to disentangling the impact of family planning from the role of other maternal and child health services that were also offered as part of the Matlab pro-

[1]We define a person to be "gainfully employed" if they report either working for a wage or maintaining a self-enterprise

gram's suite of services. Our study contributes to this ongoing discussion by providing more direct experimental evidence of the causal impact of a family planning-focused intervention.

The rest of this paper proceeds as follows. Section 2 provides the theoretical framework linking family planning to economic outcomes. Section 3 describes the family planning intervention and experimental design. Section 4 outlines our empirical strategy, and Section 5 presents the results from our analysis of intervention impact. We discuss the implications of our findings and conclude in Section 6. In an appendix, we explore potential mechanisms behind our results through a causal mediation analysis and show that the results are partially explained by changes in fertility over the short term.

2 Theoretical Framework

We expect our family planning intervention to impact women's labor market outcomes in three ways. First, improvements in birth spacing through our intervention may relax time 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. Second, even among women who would not have become pregnant without improved access to family planning services, our intervention could enable these women and couples to more effectively control their fertility and reduce uncertainty surrounding pregnancy and birth. This improved control may, in turn, encourage women to engage with the labor market by enabling them to devote more time to productive activities 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. These effects from contraceptive use may, in turn, contribute to improvements in women's labor market outcomes through improved health and human capital.

In addition, there is also the possibility that family planning improves the labor market outcomes among children who are born to mothers who receive improved access to services. Indeed, in Maggio et al. (2023), we show that children born either directly before or at the time of the intervention's rollout were less likely to be stunted after one year of exposure to the intervention and performed better on cognitive measures after two years of exposure to the intervention. If these effects continue to accrue over these children's lifetimes, then we can infer that labor market outcomes for these children may also be positively impacted as a result of improved access to family planning. Previous research on the long-term impacts of stunting would suggest that this may be the case (Hoddinott et al., 2013); however, we would require a significantly longer follow-up period on these children to infer any potential

impacts of a family planning intervention that was rolled out to their mothers on their labor market outcomes. For this reason, we focus on mechanisms related to improved birth spacing, which we confirm in Karra et al. (2022) that improved access to family planning through our intervention was associated with a 44 percent reduction in the relative risk of short birth intervals for women assigned to the intervention arm.

While there is likely significant correlation between fertility and labor market outcomes, better labor market conditions may in turn cause reductions in fertility as higher incomes increase the opportunity cost of childbearing. In response to this increased opportunity cost, women and couples may reduce their birth parity. While we cannot directly estimate this relationship in this study, we test the connection between a fertility-reducing intervention and labor market outcomes using a causal mediation analysis, which we present in the appendix. A goal of this analysis is to determine how much of the impact from the intervention can be attributed to changes in short-term fertility.

3 Study Design

Our study is based on 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 as it relates to our present analysis; a more detailed description of the study protocol and implementation of the intervention is presented elsewhere Karra and Canning (2020).

As part of the study, 2,143 women[2] who were either pregnant or immediately postpartum (within 6 months of their last live birth), between the ages of 18 and 35, and were living in Lilongwe were recruited in 2016. Following a baseline survey, women were individually randomized to either an intervention arm or control arm. Women assigned to the intervention arm received a comprehensive family planning package that consisted of private family planning counseling sessions, free transport to a high quality family planning 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 again do not observe any large or statistically significant differences across treatment arms. In our sample, 23.5 percent of women reported being employed, defined as engaging in wage-earning labor or being self-employed, at baseline, with a majority (XX percent) of these women reporting being self-employed. We find the opposite for men, 92.4 percent of whom

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

report being employed, with a majority (XX 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.

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 is Employed	0.232	0.237	0.235	-0.005
Wage Earning Employment	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 is Employed	0.920	0.928	0.924	-0.008
Wage Earning Employment	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.

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, followed by two follow-up surveys. Data collection for the first follow-up survey began in August 2017 and was completed in February 2018, and data collection for the second follow-up survey 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. However, not all 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 those presented in Karra et al. (2022).[3]

3.1 The Intervention

Women assigned to the intervention arm were offered a comprehensive, multi-component family planning package over a two year service period. Designed in coordination with local health authorities, NGOs, and communities, the intervention aimed to overcome multiple barriers to access to care. As such, the package combined family planning counseling with free transport and financial reimbursement to overcome key barriers to family planning access in Lilongwe. By offering a comprehensive package, this study aims to determine the causal effect of improved access to family planning.

As part of the intervention, women were offered up to six free, private family planning 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 family planning 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 also offered free transportation (private taxi rides) to a local private family planning clinic, the Good Health Kauma Clinic, which offered women comprehensive family planning 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 MKW (∼$25.00 USD) in financial coverage for any costs that they incurred while receiving family planning care at the Kauma Clinic, or any other clinic in Lilongwe that women chose to attend. Costs

[3]Using less detailed questions, we can 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 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, family planning for women with HIV or other health concerns, and comprehensive treatment and management of side effects.

that were eligible to be reimbursed included those related to the procurement and removal of contraceptive methods, family planning 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 family planning and reproductive health care. Though the reimbursement allowance amount was non-transferable to non-family planning 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 family planning clinic. The women were given this information during the baseline survey and were only contacted again for 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, employment, and labor income 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 addition, T_i represents the treatment assignment for women i, Y_{i0} represents the baseline value of the outcome, and ε_{it} represents the 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.

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

In our models, the coefficient β_T identifies the intent-to-treat effect of being assigned to 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 paper 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 family planning components during their first family planning visit; 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 intention to treat effect allows us to infer the population-level effect of improved access to family planning services in a context where service uptake may be only relevant and 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).

In this study, we present results on a range of outcomes related to women's labor force participation, employment, and labor supply. Our main outcome of interest is women's employment, where we consider a Woman to be employed if she reports conducting wage-earning labor or running an enterprise within the past week. For those respondents who report earning a wage, we present results on both the extensive and intensive margins of their reported labor income. We report results for both men and women for each outcome of interest. In the appendices, we also report results on the impact of our intervention on agricultural labor supply, labor force participation, unemployment, and time use.

4.2 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, and women who were immediate postpartum. 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. To this end, Section 5.2 presents a subgroup analysis based on women's baseline pregnancy status as well as baseline educational attainment and employment status.

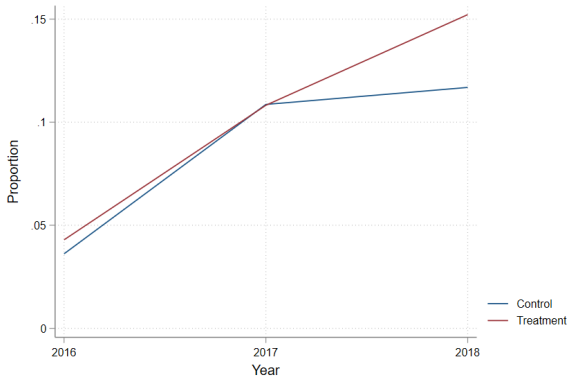
5 Results

5.1 *ITT Effects*

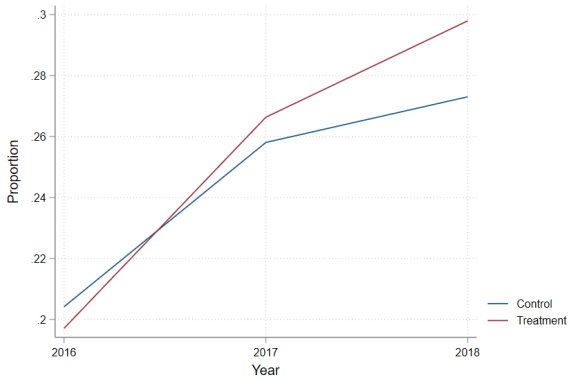
We first describe our main outcomes over time. Figure 1 presents employment rates over the study period for both men and women. For women, we observe similar trends among the treatment and control arms of our intervention during the first year of the intervention.[6] However, during the second year of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

Figure 1: Work Outcome Trends over Time

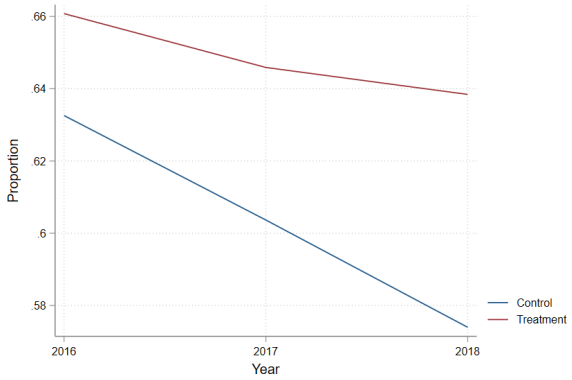
(a) Women's Wage-Earning Work



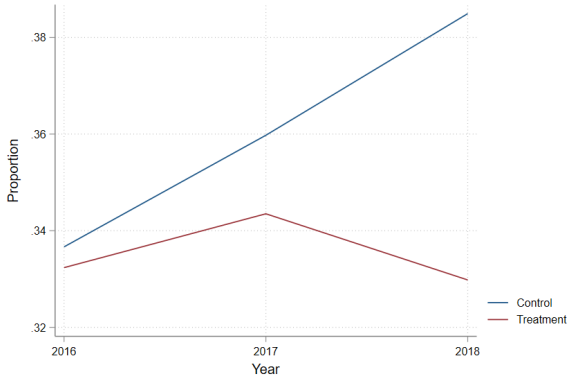
(b) Women's Self-Employment



(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.

In Table 2, we investigate these trends empirically and present the ITT results on women's

[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.

labor market outcomes at 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 5.2 p.p. more likely to be employed. This effect on employment is likely driven by increased participation in wage-earning labor; women assigned to the intervention arm are 3.3 p.p. more likely to report being engaged in wage-earning labor and are 3.8 p.p. more likely to report earning labor income at endline if they were assigned to the intervention arm. However, we note that while women in the intervention arm may be more likely to partake in wage-earning labor, they do not earn more income when compared to wage-earning women in the control arm.

While our point estimates on employment are positive, our estimates self-employment 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. In the accompanying appendices (XX), we use a causal mediation analysis to explore the mechanisms driving our results. We observe that the increase in employment is partially driven by changes in women's fertility over the short term, implying that as pregnancies are avoided or spaced, women can more effectively participate in the labor market. However, this pathway only explains roughly 16 percent of the total effect of the intervention that we observe. The remaining portions of the effect cannot be explained by changes in women's contraceptive use or health (proxied by their anemia status), suggesting an alternative underlying mechanism.

Table 2: ITT Effects on Women's Labor Outcomes

	(1)	(2)	(3)	(4)	(5)
				LN(Labor Income)	
	Formal Employment	Wage Earning Employment	Self-Employment	Extensive Margin	Intensive Margin
Unadjusted Estimate	0.0526	0.0353	0.0249	0.0382	-0.403
	(0.0241)	(0.0168)	(0.0222)	(0.0174)	(0.264)
	[1672]	[1672]	[1672]	[1514]	[198]
ANCOVA Estimate	0.0523	0.0338	0.0276	0.0387	
	(0.0237)	(0.0166)	(0.0216)	(0.0174)	
	[1668]	[1668]	[1668]	[1506]	
Fully Adjusted Estimate	0.0491	0.0324	0.0259	0.0387	-0.454
	(0.0237)	(0.0169)	(0.0217)	(0.0176)	(0.268)
	[1666]	[1666]	[1666]	[1504]	[198]

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 naive 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 sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

Our findings are promising for understanding the broader welfare implications of family planning. However, we may also expect to observe unintended negative consequences if

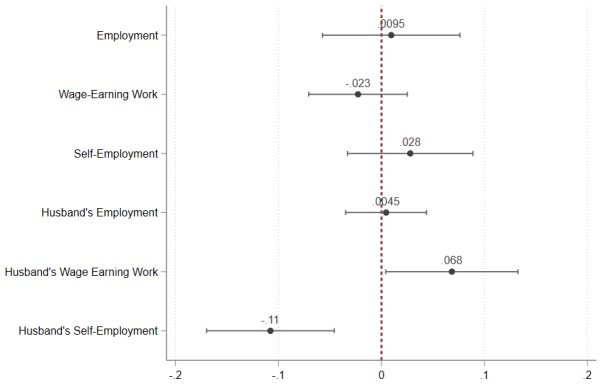
households view female and male labor supply as substitutes; if this is the case, then we may observe a commensurate reduction in male labor towards female labor as fertility falls. To explore this possibility, we document trends in male employment and labor supply in Figure 1. In contrast to women's employment, we observe a lack of trends between wage-earning employment and self employment in men. The insignificant differences in employment 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 in Table 3. Overall, we do not observe a compensatory decrease in men's labor outcomes in response to women's increased employment, suggesting an increase in the total household labor supply. As suggested by Figure 1, we find a 5.3 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 may indicate that as family planning services allow women to increase their participation in the labor market, their husbands can dedicate less of their time toward household enterprises, opting instead to participate in more formal, waged types of employment.

Table 3: ITT Effects on Men's Labor Market Outcomes

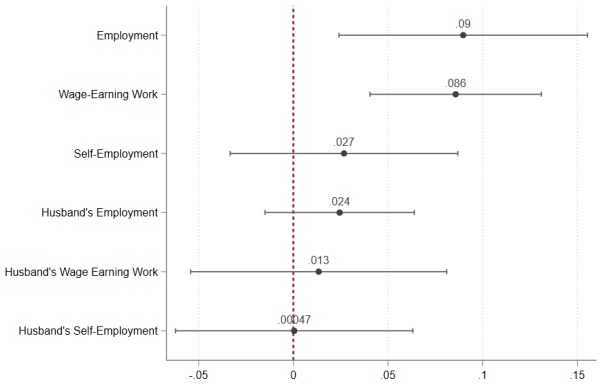
	(1)	(2)	(3)	(4)	(5)
				LN(Labor Income)	
	Formal Employment	Wage Earning Employment	Self-Employment	Extensive Margin	Intensive Margin
Unadjusted Estimate	0.0170	0.0645	-0.0551	0.0437	-0.0315
	(0.0141)	(0.0262)	(0.0257)	(0.0268)	(0.232)
	[1391]	[1391]	[1391]	[1391]	[732]
ANCOVA Estimate	0.0162	0.0441	-0.0531	0.0257	
	(0.0141)	(0.0237)	(0.0223)	(0.0254)	
	[1383]	[1383]	[1383]	[1383]	
Fully Adjusted Estimate	0.0149	0.0407	-0.0513	0.0241	0.00222
	(0.0142)	(0.0237)	(0.0223)	(0.0255)	(0.235)
	[1382]	[1382]	[1382]	[1382]	[731]

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 naive 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 wives's age, age of sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

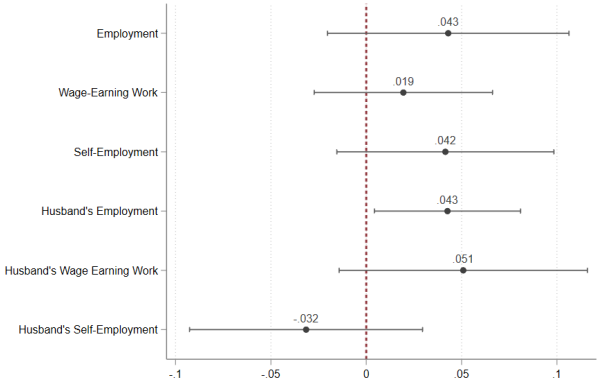
Figure 2: Coefficient Plots of Heterogeneous Effects at Study Endline



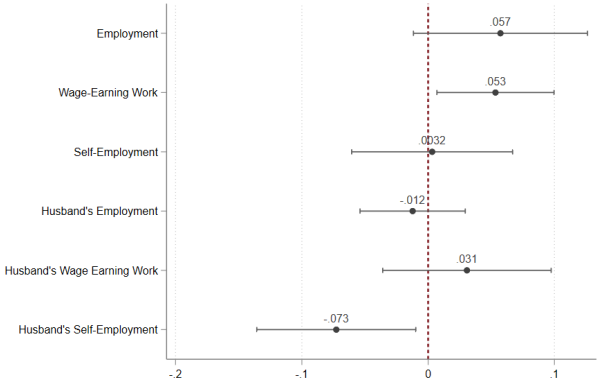
(a) Postpartum at Baseline



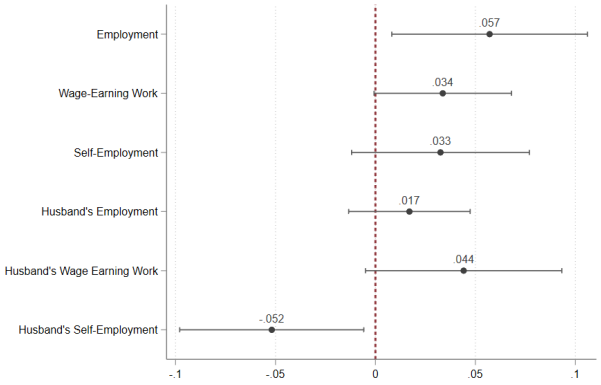
(b) Pregnant at Baseline



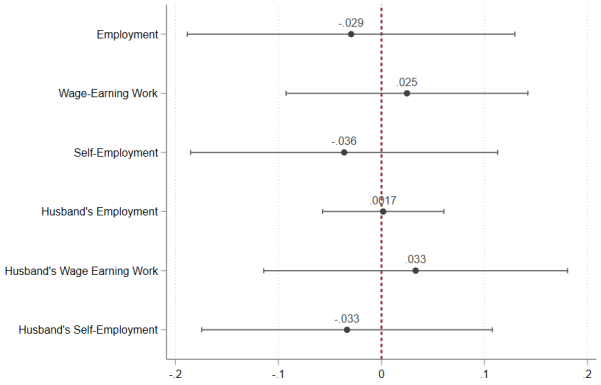
(c) No Secondary Education



(d) Secondary Education



(e) Not Employed at Baseline



(f) Employed at Baseline

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

5.2 Heterogeneous Treatment Effects

Figure 2 presents findings from the heterogeneity analysis. Specifically, we present the coefficients and confidence intervals of the adjusted ITT effects on women's and their husbands' labor outcomes by pregnancy status at baseline, women's baseline educational attainment, and employment 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 of 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. This finding suggests 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 employed at baseline largely correlate with the average treatment effects across the entire sample. This may be expected as most of the sample was not employed at baseline. However, it also may reflect the persistence of labor market activities throughout and following pregnancy, as women who are already employed were unlikely to change their behavior due to the intervention.

6 Discussion

Much of the excitement about the "demographic dividend" relies on one presumed microeconomic foundation; that reductions in fertility increase household labor supply. This paper uses experimental evidence to determine if that economic relationship holds. Our ITT estimates support this relationship and show that large labor market gains are associated with the provision of family planning services; women are 5.7 p.p. more likely to work and 3.4 p.p. more likely to participate in wage-earning labor. The gains are especially large for women who were pregnant when they began receiving the intervention. These women are 9 p.p. more likely to participate in the labor market due to the comprehensive family planning intervention. Furthermore, we do not observe deleterious effects on men's labor outcomes,

easing concerns about compensating decreases in labor market participation due to the intervention. These results provide promising evidence that the underlying assumption of the "demographic dividend" does hold in practice.

Our results contribute to a large literature linking women's fertility to their labor market outcomes (Angrist and Evans, 1996; Chun and Oh, 2002; Cruces and Galiani, 2007). While much of this evidence comes from quasi-experimental results, we can supplement this literature with evidence from a randomized control trial. We also contribute to a growing literature on the long-term impacts of family planning services. Although over a much shorter time frame and in a much different context, our results stand in opposition with the results of Barham et al. (2021) who do not find evidence of labor market improvements among women in Matlab, Bangladesh exposed to the Maternal and Child Health/Family Planning (MCH/FP) Program.

Much like our prior experimental findings (Karra et al., 2022; Maggio et al., 2023), the results presented here are a piece of a more extensive use case for family planning services. All governments and organizations have budget constraints. If cost-benefit analyses of family planning services only account for the positive effects on first-line outcomes while ignoring the positive externalities such as those presented here, they are likely to under-invest in these services. These results suggest that policymakers must consider the effect on outcomes beyond contraceptive use and fertility when considering the design of family planning programs.

References

Acharya, A., M. Blackwell, and M. Sen (2016, August). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. *American Political Science Review 110*(3), 512–529. Publisher: Cambridge University Press. 17

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, 2, 14

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, 14

Baron, R. M. and D. A. Kenny (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of*

Personality and Social Psychology 51(6), 1173–1182. Place: US Publisher: American Psychological Association. 18

Becker, G. (1960). An Economic Analysis of Fertility. NBER Chapters, National Bureau of Economic Research, Inc. 1

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. 2

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

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, 14

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, 14

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. 2

Hoddinott, J., J. R. Behrman, J. A. Maluccio, P. Melgar, A. R. Quisumbing, M. Ramirez-Zea, A. D. Stein, K. M. Yount, and R. Martorell (2013, November). Adult consequences of growth failure in early childhood. *The American Journal of Clinical Nutrition 98*(5), 1170–1178. 00150. 3

Imai, K., L. Keele, and D. Tingley (2010). A general approach to causal mediation analysis. *Psychological Methods 15*(4), 309–334. Place: US Publisher: American Psychological Association. 18

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. 17

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. 2

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. 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. 4, 5, 6, 8, 14

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

Maggio, D., M. Karra, and D. Canning (2023, November). Family Planning and Children's Human Capital: Experimental Evidence from Urban Malawi. *Working Paper*. 3, 14

Malthus, T. R., D. Winch, and P. James (1992, August). *Malthus: 'An Essay on the Principle of Population'*. Cambridge University Press. Google-Books-ID: 4EQTCHiFUBMC. 1

Pearl, J. (2014, December). Interpretation and identification of causal mediation. *Psychological Methods 19*(4), 459–481. 17

Rosenzweig, M. R. and K. I. Wolpin (1980). 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. (2009). How does Family Planning Promote Development? Evidence from a Social Experiment in Matlab, Bangladesh, 1977 – 1996. 2

A Appendix A: Women's Time Use

Table A1: 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	0.247	-0.00656	-0.120	0.0412
	(0.165)	(0.0849)	(0.155)	(0.0299)
Observations	1514	1514	1514	1514
R Squared	0.0210	0.0290	0.0467	0.00136

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

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 naive 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 why we may expect family planning services to improve women's labor market outcomes, citing the relaxation of time constraints created by the reduced burden of childcare. Within our data, we can test this hypothesis by presenting

the effects of our intervention on women's time use.[7] In Table A1, we present these results for the first-year follow-up and the study's endline. Across waves, we do not observe evidence of an effect on women's time use. At the study's endline, our estimates are quite imprecise, but their signs would imply that women spend more time at work and substitute away from both housework and child care. In practice, given the observed effects on labor outcomes, we view it as unlikely that there are truly no impacts on time allocation. Instead, we believe that we are underpowered to detect changes in allocation over a short 24-hour window.

B Appendix B: Mediation Analysis

Using this methodology, we show that direct changes in birth parity over the short term can explain roughly 16 percent of the measured effects. While this may suggest that the direct fertility effects on labor market conditions are small, it is important to note that we can only measure fertility over a very short window. There may be indirect "empowering effects" of certainty over fertility driving a potentially large portion of our results. These "empowering effects," however, are more difficult to measure. We discuss in Section 2 of the main text that our analysis throughout this paper is motivated by a theoretical framework linking family planning to labor outcomes through either reduced fertility or increases in labor productivity due to improved health. While our results are promising, testing these hypotheses would also be ideal. As such, we conduct a causal mediation analysis within this appendix to test the mechanisms underlying our results. Our approach to causal mediation analysis seeks to identify the causal mediation effect (Pearl, 2014; Imai et al., 2010; Acharya et al., 2016), given by:

$$\delta(t) = Y_i\left(t, M_i(1)\right) - Y_i\left(t, M_i(0)\right) \tag{B1}$$

where $M(t)$ is a variable affected by the treatment and lies on the causal path between the treatment and the outcome. The causal mediation effect represents the change in the potential outcome induced by changes in the mediating variable under treatment regime t. Put more simply; the causal mediation effect is the change in the outcome created by changes in the mediator in response to treatment. In the context of our theoretical framework, the causal mediating effect of fertility on labor outcomes is the amount of our treatment effect caused by changes in fertility resulting from treatment.

Fundamentally, the causal mediation effect is unidentified because we are unable to observe both $Y_i(t, M(1))$ and $Y_i(t, M(0))$. However, Imai et al. (2010) shows that the *average causal mediation effect* (ACME), $\bar{\delta}(t) = E[Y_i\left(t, M_i(1)\right) - Y_i\left(t, M_i(0)\right)]$ can be parametrically

[7] Time use is calculated using a 24 hr recall of the women's activity by the hour.

identified within a Baron and Kenny (1986) Linear Structural Equation Model (LSEM) structure. In this study, we use binary mediators and opt for a more flexible model than the LSEM. Specifically, we estimate the treatment-mediator relationship using a binary dependent variable logistic model. Our set of equations is given by:

$$M_i = \beta_0 + \beta_1 T_i + \mathbf{X}_i \beta_2 \tag{B2}$$

$$Y_i = \alpha_0 + \alpha_1 T_i + \mathbf{X}_i \alpha_2 \tag{B3}$$

$$Y_i = \gamma_0 + \gamma_1 T_1 + \gamma_2 M_i + \mathbf{X}_i \gamma_3 \tag{B4}$$

Here, Y_i is the employment status of woman i, T_i is the treatment status of woman i, and M_i is one of three mediating variables. Finally, $\mathbf{X}_i$ represents a vector of pre-treatment covariates. To identify the ACME, our system of equations must meet a sequential ignorability assumption, meaning our mediator and outcome cannot be jointly determined, conditional on $\mathbf{X}_i$ and T_i. If the sequential ignorability assumption holds, Imai et al. (2010) shows that the ACME is identified by $\alpha_1 \gamma_2$

Table B1: Mediation Analysis Results on Women's Employment

	(1)	(2)	(3)
ITT on Mediator	-0.039	0.056	-1.577
	(0.0131)	(0.0214)	(0.968)
ACME	0.007	0.002	0.0044
	[0.0019,0.0143]	[-0.0009,0.0064]	[-0.0008,0.0127]
Mediator	Short-Term Fertility	Contraceptive Use	Hemoglobin
Observations	1,600	1,666	930

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

Notes: For all columns, the unit of observation is a woman. Results presented are for a causal mediation analysis of the ITT results on employment. Woman-level controls include the total number of children who are alive, educational attainment of the woman (primary or less versus secondary and higher), age of the woman, religion (Christian versus other), tribal ethnicity (Chewa versus other), and age of sexual debut. These same controls adjust ITT estimations on the mediator. Heteroskedasticity-robust standard errors, are presented in parentheses, and bootstrapped confidence intervals are included in brackets.

In this study, we present analysis for three mediating variables in line with the theoretical framework presented in the main text. First, we test the mediating effect of short-term fertility, taking the value one if a woman had a subsequent birth event or pregnancy beyond the index birth during our study period. Acknowledging that the study window was short and unlikely to internalize all fertility effects, we also test our equations using contraceptive use as the mediator to detect any effects related to certainty over fertility. Finally, to test

biological pathways, we test the mediating effects of hemoglobin.[8]

In Table B1, we present the results of our mediation analysis. Although we observe positive point estimates for each of our estimations, we only observe statistically significant effects when mediating with short-term fertility. Here, we find an average causal mediation effect of 0.7 p.p., meaning that changes in fertility throughout our study can explain roughly 16 percent of the effects on employment. While this is only a small proportion of the total effect, it does lend credence to the idea that family planning relaxes time constraints on women by enabling them to space births or reduce their fertility.

In total, even including those point estimates where we do not have strong evidence of a mediating effect, we can only explain roughly 22 percent of the total effects we observe on employment. This builds the need for more evidence on the mechanisms underlying the causal link between employment and family planning that we have identified here.

C Appendix C: Secondary Work Results

C.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 a main outcome because it is likely not indicative of wellbeing 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. For this reason, the data on these outcomes is incomplete.

Nevertheless, in Table C1, we present the effects on these outcome for women in our 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 employment in the main text and the fact that we can rule out compensating decreases in

[8] We observe a high rate of missingness in the measurement of women's hemoglobin, meaning the sample size of this third estimation is substantially smaller than the previous two. To test the sensitivity of our results to this missingness, we calculate the mediating effects of hemoglobin using multiple imputations and do not find significant changes in the results. We opt to present the mediation analysis here using the true data.

Table C1: ITT Effects on Women's Labor Market Outcomes

	(1) Agricultural Work	(2) Unemployment	(3) Labor Force Participation
Unadjusted Estimate	-0.0000360	0.00855	0.0384
	(0.00463)	(0.0114)	(0.0254)
	[1670]	[1514]	[1514]
ANCOVA Estimate	-0.0000336	0.00827	0.0409
	(0.00465)	(0.0114)	(0.0249)
	[1665]	[1506]	[1506]
Fully Adjusted Estimate	-0.0000793	0.00850	0.0374
	(0.00472)	(0.0114)	(0.0249)
	[1663]	[1504]	[1504]

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 naive 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 wives's age, age of sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

unemployment due to the precision of our estimates in column (1), it may be 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 C1 with findings for men in Table C2. 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 C1, 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.

C.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 C3 for women and Table C4 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. Although there is some evidence of improvement in labor outcomes

Table C2: ITT Effects on Men's Labor Market Outcomes

	(1) Agricultural Work	(2) Unemployment	(3) Labor Force Participation
Unadjusted Estimate	0.00482	-0.000535	0.0112
	(0.00389)	(0.0103)	(0.00702)
	[1391]	[1391]	[1391]
ANCOVA Estimate	0.00470	-0.000618	0.00965
	(0.00387)	(0.0103)	(0.00682)
	[1383]	[1383]	[1383]
Fully Adjusted Estimate	0.00448	0.000479	0.00934
	(0.00407)	(0.0105)	(0.00677)
	[1382]	[1382]	[1382]

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 naive 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 wives's age, age of sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

among men when studying our unadjusted estimates, these estimates retain their sign but lose traditional statistical significance when controlling for baseline values of each outcome.

Table C3: ITT Effects on First-Year Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							LN(Labor Income)	
	Formal Employment	Wage Earning Employment	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 naive 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 wives's age, age of sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.

Table C4: ITT Effects on First-Year Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							LN(Labor Income)	
	Formal Employment	Wage Earning Employment	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 naive 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 wives's age, age of sexual debut, birth history, contraceptive experience and use, education, work status, religion, ethnicity, and pregnancy status. Heteroskedasticity-robust standard errors are reported.