Family Planning and the Dynamics of Fertility Preferences: Experimental Evidence from Urban Malawi*

Stefanie Pechar† Daniel Maggio‡ Mahesh Karra§

June 19, 2026

Abstract

Fertility preferences are core inputs in projections of demographic trends and models of household fertility and human capital investments. Such applications implicitly assume that preferences are stable over time, both in aggregate and within individuals. Using longitudinal data on stated fertility preferences from urban Malawi, we show that while preferences appear stable in the aggregate, there is substantial within-woman variation over time. We leverage a randomized expansion of family planning services to examine how the intervention shapes both the *level* and the *stability* of these preferences. We find that the intervention has no aggregate effect on desired family size but substantially reduces within-woman variability in preferences, which suggests more deliberate formation of fertility goals. These stability gains concentrate among women pregnant at baseline whose informational and access constraints were most relaxed by the intervention. We further find that increased preference stability is associated with higher contraceptive use. Our findings establish preference stability as a behaviorally meaningful dimension of reproductive decision-making and suggest that policy interventions can shape not only fertility behavior but the formation and consolidation of preferences themselves.

Keywords: Stability of Preferences; Fertility Preferences, Family Planning; Malawi

JEL Classifications: C93, D84, D91, J13, O15, I12, I15, I18

*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 uses 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. This project was supported by two grants from the Hewlett Foundation and the 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 IRB (protocol number IRB16-0421) and the Malawi National Health Sciences Research Committee (protocol number 16/7/1628). The findings, interpretations, and conclusions expressed in this study are entirely those of the authors. They do not 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.

†**Corresponding Author.** Technical University of Munich, Germany. stefanie.pechar@tum.de

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

§Mailman School of Public Health, Columbia University, New York, USA. mvk2109@cumc.columbia.edu

1 Introduction

Fertility preferences are central inputs in models of household fertility decision-making and demographic change. Their use, however, rests on the implicit assumption that these preferences are stable, well-defined, and predictive of future behavior.[1] A growing body of literature challenges this view, showing that fertility preferences are dynamic and can change substantially over time (Kodzi et al., 2010; Mueller et al., 2022). This work has primarily focused on how preferences shift in response to changing personal circumstances, such as relationship transitions, childbearing experiences, and economic shocks (Sennott and Yeatman, 2012; Yeatman et al., 2013; Hayford and Agadjanian, 2017). Much less is known about how the broader sexual and reproductive health environment—including policies and interventions that relax constraints on reproductive decision-making—shapes the formation and stability of fertility preferences.

Family planning (FP) programs provide a natural setting to study this question. Traditionally, such programs have aimed to help women and couples realize their existing fertility goals rather than to alter those goals directly (Freedman, 1997). Contemporary FP programs, however, increasingly extend beyond contraceptive provision to incorporate counseling and informational components that engage directly with reproductive decision-making (Ashraf et al., 2014; Dupas et al., 2025; Herrera-Almanza and Seitz McCarthy, 2025). These programmatic elements may shape not only whether women can implement their fertility preferences, but also how those preferences are formed, refined and consolidated.

This study investigates how fertility preferences evolve over time and how policy interventions can shape their stability. This question is especially relevant in sub-Saharan Africa, where persistent economic and health uncertainty, combined with high unmet need for contraception, may contribute to greater volatility in fertility preferences (JohnsonHanks, 2007). To do so, we combine descriptive evidence on within-woman preference dynamics with causal evidence from a randomized FP intervention using longitudinal data from urban Malawi. The intervention expanded contraceptive access through free FP services and transportation to clinics with structured counseling that encouraged pregnant and immediate postpartum women to reflect on birth spacing, desired family size, and longer-term reproductive goals. By simultaneously relaxing constraints and promoting deliberation, the program creates scope for effects not only on realized fertility behavior but also the preferences guiding it. We measure fertility preference levels using women's reported ideal number of children, and conceptualize stability as the within-woman variation in these reported preferences across

[1] For example, Becker (1960)'s standard model of fertility choice treats fertility preferences as fixed and well-defined primitives that households can optimize over and, in principle, realize.

survey waves. This treats observed changes in preferences as a behavioral outcome — capturing how women articulate fertility goals across repeated observations — rather than relying on self-reported uncertainty at a single point in time.

Our descriptive findings reveal a stark contrast in fertility preference dynamics: while aggregate preferences remain stable over time, there is substantial within-woman variation across survey waves. Importantly, this variation does not appear to reflect random reporting error. Instead, it is internally consistent with other measures of fertility intentions and follows predictable age and life-cycle trajectories. These underlying patterns imply that households may be making reproductive decisions around a parity target that is itself evolving over time. We therefore interpret within-woman variability in preferences as reflecting genuine revisions to forward-looking fertility goals, and treat this variability as a behavioral outcome in its own right.

Building on these descriptive patterns, we study how the intervention affects both the level and the stability of women's fertility preferences. Our empirical strategy proceeds in two stages: a difference-in-differences specification with individual and survey-wave fixed effects identifies treatment effects on preference levels, and a second-stage model on the absolute residuals from this specification identifies effects on within-woman variability. We complement this two-stage approach with a distributional comparison of net preference changes between treatment and control groups at endline. We further examine heterogeneity in these effects by reproductive status, education, and access constraints, and assess whether preference stability is itself associated with contraceptive behavior.

Our results show that improved access to family planning services can shape fertility preferences in ways that go beyond the aggregate level effects typically examined. On average, the intervention has no significant effect on desired family size in the full sample. This null effect conceals opposing movements across subgroups: desired fertility declines among women who are pregnant at baseline and increases among postpartum women, with only the latter effect reaching statistical significance. In contrast, the intervention has a clear effect on preference stability, significantly reducing within-woman variability in stated fertility preferences, with effects concentrated among women who were pregnant at the time of exposure — when deliberation around future reproductive decision-making is most salient — while preference stability among postpartum women remained largely unchanged. These patterns are robust across alternative model specifications and samples. Within the pregnant subgroup, these stability gains concentrate further among women with lower education and those living farthest from family planning facilities — the women whose informational and access constraints were most relaxed by the intervention. Increased preference stability is in turn associated with higher contraceptive use. Taken together, these findings suggest that

FP programs can facilitate the consolidation of fertility preferences without uniformly shifting fertility targets, with the timing of exposure moderating the operation of information and access channels. Preference stability emerges as a behaviorally meaningful dimension of reproductive decision-making in its own right.

Our paper relates to two main strands of research. First, we contribute to the literature on the measurement and dynamics of fertility preferences. A large body of work challenges the assumption that fertility preferences are fixed, showing that stated preferences or intentions are imperfect predictors of subsequent behavior and are frequently revised over time (Nair and Chow, 1980; Tan and Tey, 1994; Quesnel-Valle and Morgan, 2003; Hayford and Agadjanian, 2017; Cleland et al., 2020). On the intensive margin, longitudinal studies have documented frequent within-person revisions in ideal family size preferences in response to life-course events such as partnership transitions, employment changes, and reproductive experiences (Heiland et al., 2008; Liefbroer, 2009; Hayford, 2009). While early evidence comes primarily from high-income contexts, more recent work extends these findings to low-income, high-fertility settings, particularly in sub-Saharan Africa, where preferences are likewise dynamic and systematically patterned (Kodzi et al., 2010; Sennott and Yeatman, 2012; Yeatman et al., 2013; Hayford and Agadjanian, 2017). A complementary strand of this literature focuses on measuring preference uncertainty and variability, typically using survey elicitation methods such as ranking exercises or graded confidence categories (Coombs, 1974; Morgan, 1981; Barker and Buber-Ennser, 2024; Badolato et al., 2025). While informative, these approaches are largely cross-sectional and rely on individuals' ability to reliably report the strength of certainty surrounding inherently uncertain preferences, providing limited insight into how preferences evolve over time.

Second, we contribute to the literature on the impact of FP programs in high-fertility contexts. Existing studies primarily focus on behavioral outcomes, such as contraceptive use, birth spacing, and fertility, with mixed results that vary by baseline demand and constraints (Joshi and Schultz, 2013; Karra et al., 2022; Desai and Tarozzi, 2011; Dupas et al., 2025). Even when measured, fertility preferences typically appear as secondary outcomes or model inputs used to explain observed behavior or motivate causal pathways rather than studied as outcomes in their own right. One notable exception is recent work by (Herrera-Almanza and Seitz McCarthy, 2025) which evaluates the impact of an informational FP intervention on desired additional children.

Bringing these literatures together, we make three contributions. First, we conceptualize and operationalize within-woman variability in fertility preferences as a behavioral measure of preference stability, moving beyond cross-sectional and self-reported measures of uncertainty. Second, using panel data from a randomized experiment, we provide causal evidence on how

improved access to FP services shapes both the level and stability of fertility preferences over time. Third, by examining preference stability as an outcome of programmatic exposure, we show that policy interventions can influence not only fertility behavior but the persistence of preferences themselves, shifting attention from individual-level determinants toward the role of the policy environment in structuring fertility goals.

2 Background

2.1 Fertility Transition and Reproductive Preferences in Malawi

Our study takes place in Malawi, a setting characterized by a gradual but incomplete fertility transition. Total fertility has declined substantially over recent decades, from seven births per woman in the early 1980s to roughly four to five births in recent years (ICF, 2017). This decline has been accompanied by increases in female education, improvements in child survival, and expanding access to reproductive health services. At the same time, fertility remains high relative to global averages, with substantial heterogeneity across regions, socioeconomic groups, and life stages. This combination of declining fertility and continued high reproductive exposure places Malawi in an intermediate stage of the demographic transition, where fertility preferences are actively evolving and subject to ongoing reassessment.

Consistent with this view, ideal family size has declined over time, but slowly and unevenly, and large families continue to be widely valued in many settings. Survey evidence from Malawi and similar contexts shows considerable dispersion in stated fertility preferences and substantial within-woman revisions to preferences over time (Kodzi et al., 2010; Trinitapoli and Yeatman, 2018; Yeatman et al., 2013). These patterns suggest that preferences are responsive to changing circumstances rather than reflecting stable long-term targets.

This dynamic environment has important implications for family planning. In a setting where fertility goals are still being formed and revised, programs that expand access to contraception and provide reproductive counseling may affect not only the implementation of existing preferences but also how those preferences are articulated and maintained over time. Malawi therefore provides a natural setting in which to study how changes in the policy and service provision environment shape both the level and stability of fertility preferences.

2.2 Conceptual Framework

The empirical literature on fertility preferences has extensively documented that stated ideals vary across individuals and over time but offers limited guidance on how to interpret this

variation for measurement and policy. We develop a conceptual framework that organizes the sources of within-woman variability in fertility preferences, clarifies the channels through which a FP intervention may affect this variability, and motivates the empirical patterns and heterogeneity in treatment effects that we observe. Our framework draws on insights from demography, economics, and social psychology to argue that within-woman variability in stated preferences is a structured outcome that reflects ongoing processes of preference formation under uncertainty.

2.2.1 Fertility Preferences under Uncertainty

We depart from the standard view of fertility preferences as fixed parameters[2] and instead conceptualize them as constructed under uncertainty. Drawing on the cognitive-social model of Bachrach and Morgan (2013) and related work (Bhrolchin and Beaujouan, 2019), we view stated preferences as evaluative judgments that are actively constructed in response to available information, constraints, and the reproductive context at the time of elicitation.[3]

Formally, let a woman i's reported ideal family size at time t be given by

$$\text{IFS}_{it} = f_i(\Omega_{it},\, C_{it},\, R_{it}) + \varepsilon_{it}, \tag{1}$$

where Ω_{it} denotes her information set at time t (encompassing knowledge about contraceptive methods, health risks, partner preferences, and economic prospects, among others), C_{it} captures the perceived constraints on her reproductive choices (including contraceptive access, financial resources, spousal bargaining power, etc.), and R_{it} represents her current reproductive state (e.g. pregnant, immediate postpartum). The function $f_i(\cdot)$ maps these inputs into a latent preferred family size that is specific to each woman. Crucially, none of the arguments of $f_i(\cdot)$ need to be time-invariant: as information accumulates, constraints shift and reproductive circumstances evolve, and constructed preferences update in turn. To distinguish genuine uncertainty in the constructed preference from statistical noise, we further decompose the residuals ε_{it} as

$$\varepsilon_{it} = u(\Omega_{it},\, C_{it},\, R_{it}) \cdot \eta_{it} + v_{it}, \tag{2}$$

[2]Specifically, standard demographic and economic models of fertility typically treat preferences as given, either as stable parameters that enter the household's optimization problem (Becker, 1960) or as fixed ideals against which realized fertility is compared (Bongaarts, 1990). Under these views, within-woman changes in stated preferences primarily reflect measurement error or transient reporting noise, and the analytical focus falls on the gap between preferences and outcomes rather than on preferences themselves.

[3]Under this view, reported ideal family size reflects not the recall of a predetermined number but a synthesis - consciously or otherwise - of a woman's current understanding of economic constraints, health considerations, partnership dynamics, social norms, and prior childbearing experiences into a single summary response (Bhrolchin and Beaujouan, 2019).

where $u(\cdot)$ captures uncertainty in the latent preferred family size and v_{it} is statistical noise in ideal family size that may arise from measurement error, survey effects (e.g. priming), and other cognitive biases. The function $u(\cdot)$ takes the same arguments as $f_i(\cdot)$ but is not indexed by woman: while the partial derivatives of f_i with respect to (Ω, C, R) may differ in direction across women, we assume that changes in these arguments move uncertainty uniformly. This structure allows a change in any argument to affect the precision of preferences without necessarily shifting aggregate ideal family size. Observed within-woman variability in stated preferences therefore need not be interpreted solely as noise; it may also reflect structural change in the determinants of preferences.

2.2.2 Sources of Within-Woman Variability

Within this framework, three distinct mechanisms generate within-woman variability in stated fertility preferences.

First, *updating the information* leads to a revision in fertility preferences as women acquire new knowledge about the costs and benefits of childbirth. Learning about one's own fecundity, the health risks from closely spaced births, the economic demands of additional children, or the preferences of a partner can shift the basis on which fertility preferences are formed. In environments characterized by high uncertainty, such as low-income settings with volatile economic and health conditions, preferences may be revised frequently as women respond to new information over time (Trinitapoli and Yeatman, 2018; Yeatman et al., 2013).

Second, *deliberation* affects how well-defined preferences are at any given point in time. Research on attitude strength in social psychology demonstrates that attitudes that have been actively elaborated - through reflection, discussion, or comparison against alternatives - tend to be more stable and resistant to change than those that are weakly formed (Howe and Krosnick, 2017). Applied to the fertility context, a woman who has carefully considered her desired family size and weighed relevant trade-offs is more likely to report consistent preferences over time than a woman for whom the question evokes only a loosely held intuition. Importantly, this channel does not change the underlying direction of fertility preferences but rather operates through a degree of cognitive engagement by consolidating the preferences that emerge from the deliberative process.

Third, *implementability* of preferences shapes the precision with which they are articulated. When structural barriers, such as limited access to contraception, financial constraints, or informational gaps, prevent women from acting on their reproductive goals, the practical value of holding a precise fertility target is limited. In such contexts, preferences may remain vague or loosely specified. Conversely, when these constraints are relaxed and women

acquire the means to regulate their fertility more effectively, preferences over family size become more consequential, increasing the incentive to resolve ambiguity and commit to more stable targets. This mechanism implies that preference variability is partly endogenous to the policy environment: the same woman may report more variable preferences when she lacks the means to implement them, and more stable preferences when effective means to achieve her desired family size are within reach.

2.2.3 Intervention Channels

FP interventions can engage all three mechanisms, creating scope for effects on preference stability that are distinct from effects on preference levels. Through counseling components, such interventions can directly expand women's information sets, thereby reducing variability arising from uncertainty about reproductive health and contraceptive options. When counseling is repeated and individualized, it may extend beyond information provision. By prompting women to reflect on birth spacing, desired family size, and longer-term reproductive plans, FP interventions may foster active deliberation associated with preference crystallization, providing repeated opportunities for women to construct, evaluate, and refine their fertility preferences in a structured setting.

Through access components that reduce the financial and logistical barriers to obtaining contraception and improve the quality of service delivery, FP interventions can relax constraints on fertility regulation that may otherwise discourage the formation of well-defined preferences. When contraception is reliably available and affordable, fertility goals become implementable, raising the value clear and consistent targets. Together, improved access and structured deliberation may reduce preference variability through multiple reinforcing channels: better information narrows the scope for updating, deliberation consolidates preferences, and implementability increases the returns to preference precision.

This framework implies that FP interventions may reduce within-woman variability in fertility preferences even in the absence of large average effects. If the primary channels operate through deliberation and crystallization rather than through directional persuasion, such interventions would be expected to reduce the dispersion of preferences around individual trajectories without shifting their central tendency. This is consistent with interpreting FP programs as improving the formation and consistency of reproductive preferences, rather than altering their level.

2.2.4 Interpreting Preference Stability

Our interpretation of preference stability has important welfare implications. A reduction in within-woman variability may reflect improved preference formation if instability primarily arises from uncertainty, limited information, or incomplete deliberation. In this case, greater stability indicates that women are better able to articulate coherent and well-defined fertility goals, which may facilitate alignment between preferences and realized outcomes. However, stability is not unambiguously welfare-improving. If observed variability instead reflects flexibility or responsiveness to changing circumstances, such as shifts in economic conditions, health risks, or intra-household bargaining (the very conditions that women in our study setting may face), then some degree of instability may be optimal. In such contexts, reduced variability could reflect a loss of adaptability rather than improved preference formation.

Our empirical interpretation treats reductions in variability as indicative of preference consolidation under the assumption that instability largely reflects uncertainty rather than optimal responsiveness. Consistent with this interpretation, we examine whether greater stability is associated with outcomes such as improved alignment between stated preferences and realized fertility behavior, which would suggest that increased stability reflects enhanced reproductive agency, as opposed to constrained choice.

2.2.5 Heterogeneity by Reproductive Status

The channels identified above are unlikely to operate with equal effect across women. A woman's reproductive status at the time of exposure plausibly conditions her responsiveness to the intervention, generating heterogeneous effects on preference stability.

Women who are pregnant at the time of intervention exposure are actively contemplating family size, birth spacing, and future reproductive intentions as they prepare for an impending birth. In this state, the cognitive preconditions for deliberation are already in place, allowing for the intervention's counseling component to directly engage with an active decision-making process. Research on attitude formation suggests that elaborative processing is most effective when the attitude object is salient and personally relevant (Bachrach and Morgan, 2013). Pregnancy exactly provides this condition, enabling the intervention to structure and deepen a reflective process that is already underway.

By contrast, women in the immediate postpartum period face competing demands related to infant care, physical recovery, and the logistics of managing a household with a newborn. While these women receive the same intervention components, forward-looking fertility planning may be less salient in this period. The intervention may still improve contraceptive access and provide useful information as prior analyses have shown (Karra et al.,

2022) but the consolidation of fertility preferences requires sustained cognitive engagement with reproductive goals, which the postpartum context may not readily support.

2.2.6 Measurement

Finally, our framework also informs how preference stability should be empirically measured. A common approach in the demographic literature measures the strength of fertility preferences through survey-based elicitation of subjective certainty, asking respondents how sure they are about their stated ideals or using ranking methods to infer preference strength at a single point in time (Coombs, 1974; Morgan, 1981; Barker and Buber-Ennser, 2024; Badolato et al., 2025). These measures primarily capture the deliberation dimension of our framework: preferences that have been more carefully considered will tend to be reported with greater confidence. However, they are largely silent on the other sources of variability. A woman may report high certainty about her current preference and still revise it in response to new information or changing constraints. Moreover, because certainty is assessed at the same moment as the preference itself, it is subject to the same contextual factors - information, constraints, and salience - rendering it endogenous to the preference formation process it is intended to validate.

By contrast, observed within-woman variability captures the combined effect of all three sources of instability over time, without requiring respondents to anticipate their own future consistency. It does not rely on introspective reports and directly measures whether preferences persist or change across repeated observations. We therefore treat within-woman variation in stated fertility preferences as a behavioral measure of realized preference stability, complementing rather than replacing certainty-based approaches, and study the conditions under which this stability responds to policy exposure.

2.2.7 Testable Predictions

Taken together, our framework generates several testable predictions. First, aggregate fertility preferences may appear stable even when individuals revise their preferences over time. Second, within-woman variability should exhibit structured patterns, reflecting life-cycle dynamics rather than random noise. Third, interventions that improve information, promote deliberation, and relax constraints are expected to reduce within-woman variability in preferences. Finally, these effects are likely to be heterogeneous across reproductive status, with stronger effects among women for whom reproductive decision-making is most salient.

3 Data and Experimental Design

3.1 Data and Study Setting

We use longitudinal data from the Malawi Family Planning Study (MFPS), a randomized controlled trial conducted in Lilongwe, Malawi, between September 2016 and February 2019.[4] MFPS evaluates a multicomponent FP intervention designed to address key barriers to contraceptive access, including knowledge gaps and constraints related to distance, service availability, and cost. The study follows a cohort of married women aged 18 to 35 who were either pregnant or within six months postpartum at baseline, a period early in the reproductive life course when fertility preferences are likely still forming. A total of 2,143 women were enrolled at baseline in 2016 and interviewed annually over three survey waves.

The MFPS data include detailed socioeconomic and demographic information, reproductive histories, contraceptive use, and a full set of conventional fertility preference measures, including whether a woman wants another child, the preferred timing of the next birth, and her ideal family size. In this study, we focus primarily on ideal family size (IFS), elicited consistently across waves using the following question, translated into Chichewa: *"If you could [go back to the time you did not have any children and could] choose exactly the number of children to have in your whole life, how many would that be?"* Since women in our sample are in the earlier stages of their reproductive lives and are unlikely to have exceeded their preferred fertility at baseline, we interpret IFS as desired lifetime fertility and use the terms IFS and desired family size interchangeably.

Leveraging the panel structure, we examine two complementary dimensions of fertility preferences: the level of desired family size and within-woman variation over time, which we interpret as a measure of preference stability.

3.2 Randomization and Intervention

A total of 2,140 women were enrolled in the study in 2016. Eligibility criteria required that participants were married, currently pregnant or within six months postpartum at baseline, between ages 18 and 35, a permanent resident of Lilongwe, and were not sterilized or had undergone a hysterectomy. When multiple eligible women were present within a household, the youngest was selected to participate. To minimize potential spillovers between treatment and control groups, recruitment ensured that enrolled women resided at least five households apart.

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

Following the baseline survey, women were randomly assigned to the treatment and control arm using stratified covariate-balanced randomization. Stratification was based on key baseline characteristics, including parity, current contraceptive use, age at sexual debut, educational attainment, work status, and residential neighborhood. The final allocation assigned 1,026 women to the treatment group and 1,113 to the control group. Table A1 presents key descriptive statistics for women in our study sample. Treatment and control groups are well balanced across observable baseline characteristics, with only minor differences in prior contraceptive use, and joint significance tests fail to reject systematic differences across arms at baseline. Additional details on the randomization protocol are provided in Karra et al. (2022).

The intervention was a two-year multicomponent FP program designed to address informational and access-related barriers to contraceptive use. The intervention was comprised of three main components. First, women in the treatment group received up to six individualized counseling visits delivered by trained family planning counselors. These sessions provided comprehensive information on a full range of contraceptive methods, their potential side effects, and the health benefits of birth spacing, with the aim of supporting informed decision-making over time. Second, treatment women were offered free transportation to the Good Health Kauma Clinic, a high-quality provider offering a full range of FP services with reliable supply and short waiting times. Transportation was arranged through a private taxi driver, and women who utilized the service were accompanied by a female field manager in the taxi as a means to mitigate concerns about safety or stigma. Finally, all costs incurred for FP services received at the Kauma Clinic, including costs associated with contraceptive methods, consultations, laboratory tests, and examination fees, were fully covered for the duration of the program.

Women assigned to the control arm received publicly available information on FP methods and on nearby clinics and were otherwise not recontacted until follow-up.

3.3 Previous MFPS Findings

This study is part of a broader set of analyses that present findings from the Malawi Family Planning Study. Previous studies present the main findings on pre-specified primary and secondary outcomes, including contraceptive use, birth spacing, and child health (Karra et al., 2022; Maggio et al., 2024). Overall, the intervention was effective: it increased postpartum contraceptive uptake, reduced the likelihood of closely spaced and high-risk births, and improved child health outcomes. Two years after exposure, women in the treatment group were 5.9 percentage points (p.p.) more likely to use any contraceptive method, with

larger effects (7.2 p.p.) among women who were immediately postpartum. These gains were primarily driven by the increased adoption of long-acting reversible methods, including a 4.6 p.p. rise in implant use. Crucially, the intervention also affected fertility outcomes, reducing the likelihood of a subsequent pregnancy within 24 months of the baseline birth by 43 percent.

These results indicate that the intervention not only improved access to contraception but also enabled women to exercise greater control over the timing and spacing of pregnancies. Such improvements in the ability to plan and implement reproductive decisions have important implications for the formation and evolution of fertility preferences. When structural, informational, or logistical barriers are relaxed, preferences may become more stable because intentions are more readily achievable. At the same time, increased agency may also prompt individuals to reassess and refine their desired family size. The intervention therefore provides a well-suited empirical context to examine how reductions in uncertainty and constraints shape both levels and stability of fertility preferences over time.

4 Stylized Facts on Fertility Preference Dynamics

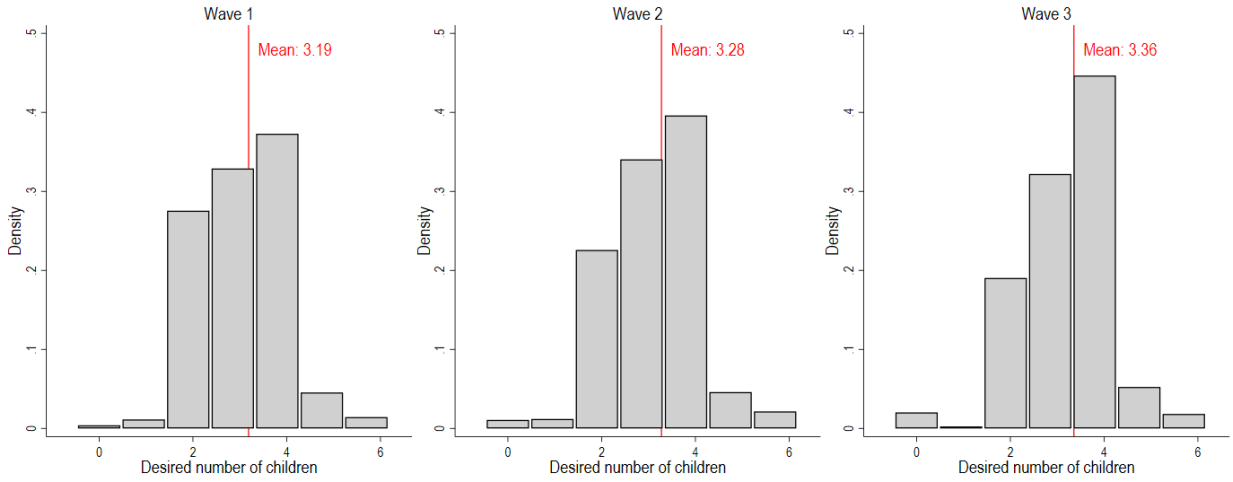
We first document patterns of fertility preference dynamics in our study setting to establish a foundation for analyzing both their level and stability. This descriptive analysis allows us to determine whether changes in reported preferences reflect meaningful behavioral processes, providing a basis for treating within-woman variability as an analytically valid outcome that captures how preferences evolve over time. We summarize these patterns through four stylized facts that characterize the core features of fertility preference dynamics in our context.

Stylized Fact 1: Fertility preferences remain relatively stable in the aggregate

Across survey waves, women in our sample report remarkably similar distributions of desired family size. In each wave, approximately 90 percent of women state a desired family size between two and four children, and the mean changes only marginally over time (Figure 1). From a population perspective, these patterns suggest a high degree of stability in fertility preferences.

At the same time, the distributions reveal small but systematic shifts across adjacent categories, notably between three and four children, while a small share of women continues to report higher desired family sizes. These patterns suggest that aggregate stability may mask offsetting movements within the distribution, reflecting revisions in individual preferences that counteract one another when summarized at the population level.

Figure 1: Desired Number of Children across Waves



Notes: The figure presents the distribution of women's reported desired number of children across survey waves. The red vertical line denotes the sample mean.

Stylized Fact 2: Substantial within-woman variability in fertility preferences

While desired family size is stable in the aggregate, it exhibits significant variation when examined longitudinally at the individual level. Figure 2 traces women's reported desired number of children across survey waves and reveals considerable movement between categories over time, indicating that many women revise their stated fertility preferences rather than maintaining a fixed target.

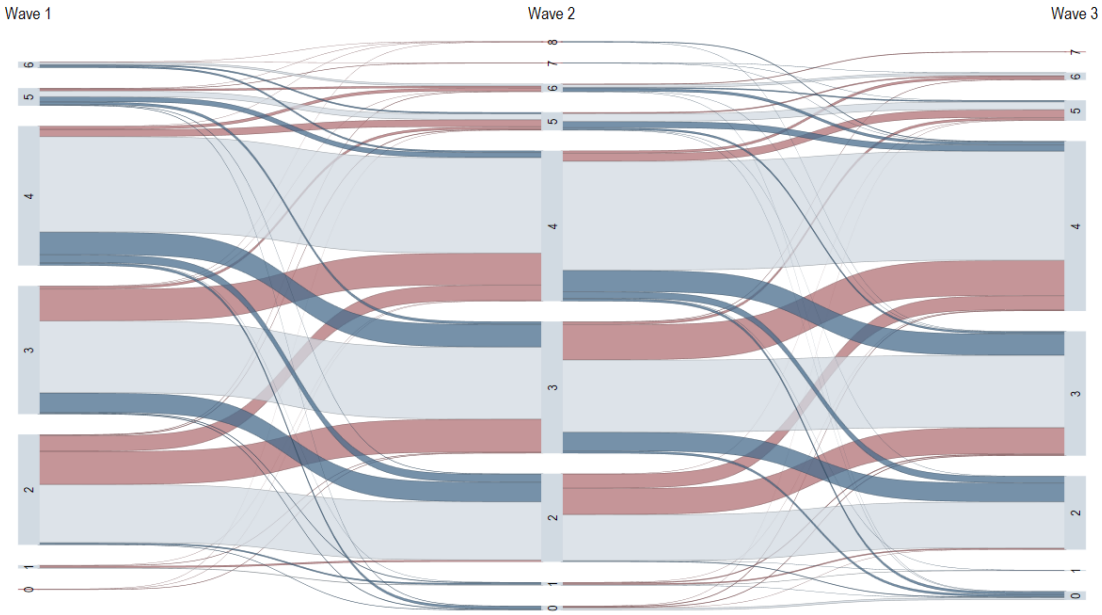
These dynamics are summarized in Table A2. Nearly 60 percent of women revise their desired family size at least once during the study period, with similar rates across age groups and baseline parity. Among those who revise their preferences, most do so once, but a substantial share (38 percent) revise twice, indicating repeated reassessment of fertility goals over a relatively short time horizon. Revisions occur in both directions: increases in desired family size are more common overall, while decreases are relatively more frequent among older women and those with higher baseline parity.

Preference revisions are often non-monotonic. Nearly one quarter of women who revise their preferences eventually revert to their baseline preference, while others revise in opposite directions across waves. Figure A1 shows that most adjustments are modest – typically by one child – but a non-trivial share of women (11 percent) report larger revisions of two or more children. These patterns indicate that within-woman variability reflects both marginal adjustments as well as more substantial re-evaluations of desired fertility.

Finally, the direction of revisions is systematically related to womens baseline fertility preferences. Women with baseline preferences below the national total fertility rate are

more likely to revise upward, while those whose baseline preferences are close to or above the national fertility level are more likely to revise downward (Figure A2). This asymmetry suggests that individual-level revisions are not random, but instead reflect movement of preferences toward context-specific reference points.

Figure 2: Within-Woman Variability in Fertility Preferences across Waves



Notes: The Sankey diagram presents individual-level transitions in ideal family size across survey waves. Flow widths are proportional to the number of women making each transition, nodes correspond to discrete ideal family size values at each wave, and colors indicate whether preferences increased, decreased, or remained stable between waves.

Stylized Fact 3: Revisions in fertility preferences are internally coherent and aligned with other fertility measures

Despite substantial within-woman variability in desired family size, fertility preferences remain internally coherent when compared to other fertility-related measures. To assess this, we examine whether womens stated fertility intentions align with the gap between their reported number of living children and their contemporaneous desired family size, specifically noting whether women below their desired family size report wanting another child, and whether those who have reached or exceeded their desired family size report not wanting additional children.

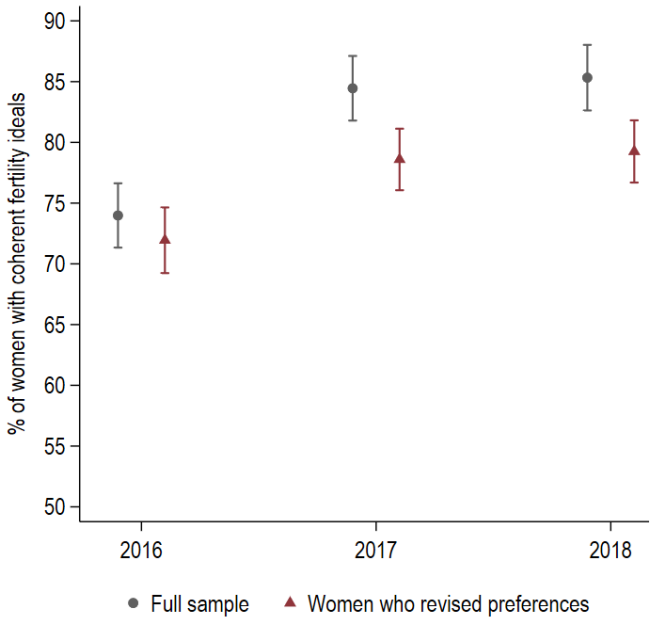
Figure 3 shows that, across survey waves, a large majority of women report fertility

intentions that are consistent with this preference-parity gap. This alignment remains high over time and is only modestly lower among women who revise their preferences compared to the full sample.

Importantly, even among women who revise their desired family size, and therefore exhibit the greatest degree of variability, approximately three quarters consistently report fertility intentions that remain coherent with their stated preferences and realized fertility. The similarity in patterns across groups indicates that revisions in desired family size are accompanied by corresponding adjustments in related measures, rather than generating contradictory responses.

Together, these descriptive patterns indicate that temporal variability in fertility preferences does not reflect incoherent reporting. Instead, fertility preferences and related measures move together in a structured manner, suggesting that preference revisions reflect ongoing reassessment rather than noise.

Figure 3: Within-Woman Consistency in Fertility Preferences



Notes: The figure reports the share of women whose fertility intentions are consistent with their reported desired family size and current parity. Consistency is defined as women with fewer living children than their desired family size reporting that they want another child, and women who have reached or exceeded their desired family size reporting that they do not want additional children. Results are shown for the full sample and for women who revised their desired family size between survey waves. Points denote sample means, and error bars represent 95% confidence intervals.

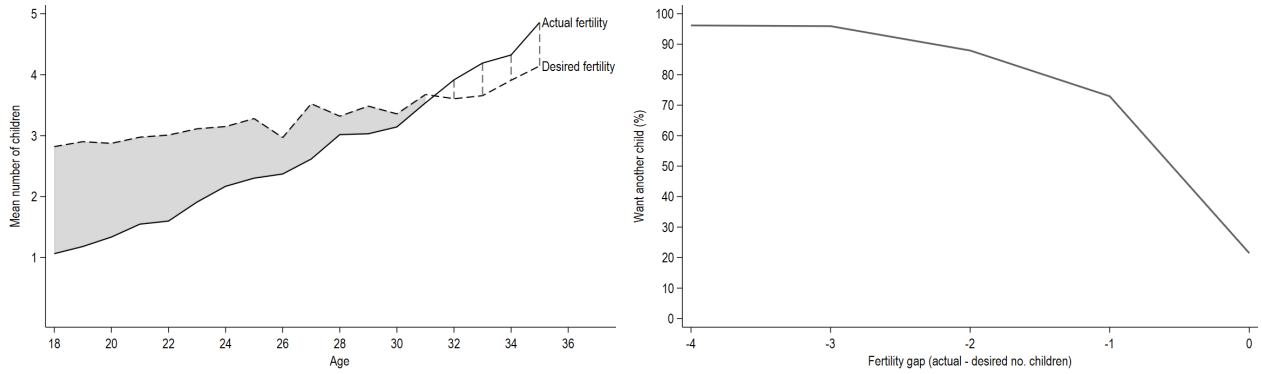
Stylized fact 4: Fertility preferences evolve predictably with age and realized fertility

Fertility preferences and childbearing intentions evolve in a structured, life-cycle consistent manner. Panel A of Figure 4 shows that reported desired family size gradually increases with age, while the gap between actual and desired fertility narrows, converging around age 32. Beyond this point, realized fertility begins to exceed desired levels, indicating that completed fertility starts to surpass earlier stated ideals.

Panel B demonstrates that this narrowing gap strongly predicts childbearing intentions. Women whose number of living children approaches their desired family size are substantially less likely to report wanting another child, while those below their target remain far more likely to express a desire for additional children. This relationship holds consistently across age groups, indicating that fertility intentions respond predictably to progress toward individual fertility goals.

Taken together, these findings indicate that temporal changes in fertility preferences and intentions follow a clear life-cycle trajectory. Desired family size and childbearing intentions evolve alongside realized fertility, reflecting structured and systematic adjustment rather than random variation.

Figure 4: Evolution of Fertility Preferences by age and realized Fertility



a) Desired and actual number of children by age b) Probability of wanting another child

Notes: Panel A displays the evolution of mean actual (solid line) and desired (dashed line) number of children by age. Panel B shows the probability of wanting another child as a function of the gap between actual and desired fertility (defined as the number of living children minus desired family size). Negative values indicate that actual fertility falls below the desired level.

5 Empirical Strategy

To estimate the causal effect of the FP intervention on both the *level* and *stability* of women's fertility preferences, we exploit the randomized design of the study and implement a two-stage empirical strategy.

In the first stage, we estimate the intention-to-treat (ITT) effect of the FP intervention on fertility preference *levels*. Exploiting the random treatment assignment, we estimate OLS regressions that control for baseline characteristics, including age, number of living children, age at sexual debut, contraceptive use, educational attainment, work status, religion, and ethnicity, as well as neighborhood fixed effects and baseline fertility preferences. In parallel, we estimate a Difference-in-Differences (DiD) specification with individual and survey-wave fixed effects, whose residuals serve as the input to our second-stage analysis of preference stability:

$$Y_{it} = \alpha_i + \gamma_t + \beta(T_i \times P_t) + \varepsilon_{it}, \tag{3}$$

where Y_{it} denotes the reported desired number of children for woman i at wave t, α_i are individual fixed effects, γ_t are survey-wave fixed effects, and ε_{it} is the error term. The interaction $T_i \times P_t$ captures exposure to the intervention. This specification leverages within-woman variation over time while controlling for period-specific shocks to estimate the treatment effect of the intervention on fertility preference levels.

In the second stage, we examine within-woman instability in fertility preferences. Following prior work that has developed approaches for estimating longitudinal variation in fertility preferences (Culpepper, 2010; Yeatman et al., 2013), we extract the residuals $\hat{e}_{it}$ from Equation 3 and define instability as their absolute value. Intuitively, these residuals capture the extent to which a woman's reported fertility preference in a given survey wave deviates from her predicted trajectory, after accounting for individual fixed effects, time effects, and treatment exposure. Taking the absolute value focuses on the magnitude of deviations, providing a measure of volatility over time.

We then estimate the effect of the intervention on this measure of instability using a Gamma generalized linear model with a log link to account for the distributional properties of the outcome:[5]

$$|\hat{\varepsilon}_{it}| \sim \text{Gamma}(\mu_{it}, \theta), \qquad \log(\mu_{it}) = \theta_0 + \theta_1 T_i. \tag{4}$$

The coefficient θ_1 captures the average difference in within-woman variability between

[5]We depart from standard OLS because the absolute deviations are non-negative, right-skewed, and heteroskedastic by construction. A Gamma specification accommodates these features by allowing the variance to proportionally increase with the mean, while the log link ensures non-negative predicted values and facilitates interpretation in proportional terms.

treatment and control groups, with negative estimates indicating greater stability among treated women. Since the residuals are estimated quantities from the first stage, using them directly as outcomes in the second stage discards the sampling variation they carry, resulting in standard errors with incorrect coverage properties. We therefore implement a bootstrap procedure that re-estimates both stages jointly across 500 replications, yielding standard errors with valid coverage. The two-stage approach allows us to separate treatment effects on the level of fertility preferences (stage 1) from effects on their stability (stage 2).

To test the channels identified in our conceptual framework (Section 2.2), we estimate Equation 4 separately on subgroups defined to isolate the women for whom each channel is theorized to operate most strongly. For the deliberation channel, under which pregnancy primes reflective engagement with the intervention's counseling content, we split by baseline reproductive status (pregnant vs. postpartum). For the information channel, under which counseling expands women's reproductive information sets, we split by highest completed education (at most primary vs. secondary or higher), where the marginal informational value of counseling is greatest among lower-educated women. For the access channel, under which the intervention's transport and cost-relaxation components ease practical barriers to contraceptive use, we split by neighborhood median distance to the nearest family planning facility.

We complement this two-stage approach with a simpler distributional comparison between groups of overall changes in fertility preferences during the study period. For each woman i, we compute the net change in desired family size between baseline and endline:

$$\Delta_i = y_{i,2018} - y_{i,2016}. \tag{5}$$

We then compare the variance of these net changes, Δ_i, across treatment and control groups:

$$H_1 : \ \mathrm{Var}(\Delta_i \mid T_i = 1) < \mathrm{Var}(\Delta_i \mid T_i = 0)$$

Under random treatment assignment, differences in dispersion can be attributed to the intervention, providing a complementary test of whether the program reduced or amplified overall shifts in preferences between baseline and endline.

6 Results

6.1 Intervention Effects on Preference Levels

Table 1 presents the estimated effects of the FP intervention on fertility preference *levels*.

Columns 1–3 report results for the full sample, while Columns 4–6 and 7–9 present estimates separately for women who were pregnant and postpartum at baseline, respectively. Across columns, specifications range from simple OLS estimates (the first column) to models with baseline controls and neighborhood fixed effects (the second column), and finally a DiD estimate with individual and survey-wave fixed effects from Equation 3 (the third column).

Across the full sample, we find no evidence of an intervention effect on fertility preference levels. Estimates are small and statistically insignificant across specifications, indicating that the program did not meaningfully shift women's stated fertility ideals, on average. However, subgroup results reveal substantial heterogeneity by reproductive status. Among women who were pregnant at baseline (Columns 4–6), the DiD specification, which leverages within-woman changes across all survey waves, yields a statistically significant decrease of approximately 0.10 children (a 3 percent decline relative to the control mean), while OLS estimates are negative but imprecise. In contrast, among women who were postpartum at baseline (Columns 7–9), the intervention significantly increases desired family size by 0.13 to 0.15 children (roughly 4 percent relative to the control mean), with effects consistently significant across specifications. Although these effects are modest in magnitude, they operate in opposite directions across subgroups and largely offset each other in the aggregate, thereby contributing to the null result in the full sample.

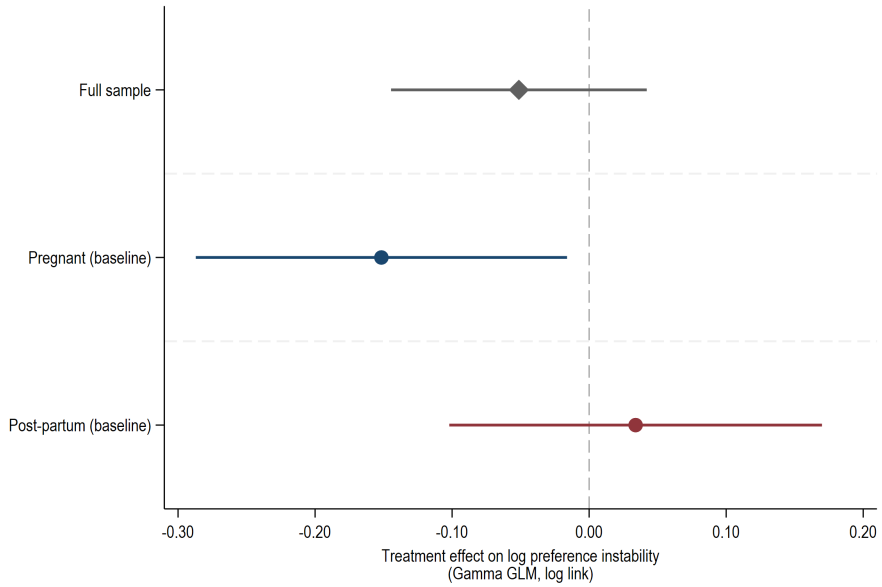
Since not all women are observed in all three survey waves, the DiD estimates includes women with incomplete panel histories. This attrition may introduce compositional differences across waves that could bias estimates, particularly if it is correlated with treatment status or fertility preferences. While Table A6 shows that attrition is unrelated to both dimensions, we nevertheless re-estimate all specifications using a balanced panel of women observed in all three waves. These results, which are reported in Table A3, are broadly consistent with the main findings. The marginally significant DiD effect among pregnant women in the main specification is no longer statistically significant in the balanced sample, which is unsurprising given its relative imprecision. In contrast, a robust effect on fertility preference levels remains for postpartum women.

Table 1: Effects on Fertility Preference Levels

	Full sample			Pregnant (baseline)			Postpartum (baseline)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.055	0.049		-0.039	-0.037		0.145**	0.135**	
	(0.047)	(0.046)		(0.066)	(0.065)		(0.067)	(0.065)	
Treatment × Post			0.012			-0.099*			0.127**
			(0.042)			(0.056)			(0.062)
Baseline Controls		√			√			√	
Neighborhood FE		√			√			√	
Woman FE			√			√			√
Year FE			√			√			√
Control mean	3.32	3.32	3.32	3.32	3.32	3.32	3.33	3.33	3.33
Number of observations	1664	1662	5276	839	838	2680	825	824	2596

Notes: The table reports the effects of the FP intervention on womens fertility preference levels for the full sample and for women who were pregnant or postpartum at baseline. Columns correspond to alternative specifications. Columns (1), (4), and (7) report simple OLS estimates. Columns (2), (5), and (8) report OLS estimates with baseline controls, including age, age at sexual debut, number of living children, contraceptive use, educational attainment, employment status, religion, ethnicity, neighborhood fixed effects, and baseline fertility preference levels. Columns (3), (6), and (9) report DiD estimates with individual and survey-wave fixed effects, with standard errors clustered at the individual level.
* p < 0.1, ** p < 0.05, *** p < 0.01

Figure 5: Effects on Fertility Preference Stability



Notes: The figure plots treatment coefficients from the second-stage Gamma GLM (Equation 4), where the outcome is the absolute value of first-stage residuals. Coefficients represent the proportional change in within-woman variability associated with the intervention, shown for the full sample and separately by baseline reproductive status. Confidence intervals are based on bootstrapped standard errors. The dashed line at zero indicates no treatment effect.

6.2 Intervention Effects on Preference Stability

Figure 5 summarizes the effects of the intervention on preference *stability* using the residual-based measure from Equation 4. Because the outcome captures within-woman variability in stated fertility preferences after accounting for time-invariant characteristics and shifts in average preference levels, negative estimates indicate *greater* preference stability that is, less unexplained variability among treated women relative to controls.

In the full sample, the estimated effect is negative but small and not statistically significant. However, the subgroup estimates reveal clear heterogeneity by reproductive status. Among women who were pregnant at baseline, the intervention significantly reduces preference instability. The estimated effect from the Gamma GLM specification implies a reduction of approximately 14 percent of within-woman variability, indicating that treated pregnant women exhibited roughly one-seventh less unexplained fluctuation in stated fertility preferences relative to controls over time.

In contrast, there is no evidence of an effect among women who were postpartum at baseline: estimates are close to zero and statistically insignificant across specifications. These patterns are robust across alternative specifications, including OLS, and across balanced and unbalanced samples (Table A4). While effect sizes differ across specifications, the qualitative result is consistent: a significant reduction in preference instability emerges only for pregnant women, with no comparable effect for postpartum women.

We corroborate these findings using a complementary approach that compares the dispersion of net changes in fertility preferences between treatment and control women at endline using a standard deviation (SD) test.[6] This distributional test likewise shows a significant reduction in variability for the full sample, which is driven by pregnant women at baseline. Consistent with the residual-variance results, no meaningful differences on stability are observed for postpartum women. The consistency in the results across approaches underscores the conclusion that the intervention enhanced preference stability specifically among women who were exposed during pregnancy, while leaving postpartum women largely unaffected.

Taken together, the results on preference levels and stability suggest that women's responsiveness to the intervention is strongly modified by their reproductive status. Among pregnant women, the intervention not only reduced desired fertility but also produced more stable preference trajectories.

[6] Unlike the residual-variance model, this approach only considers endline observations rather than the full temporal trajectory.

Table 2: Net Changes in Fertility Preferences by Treatment Status

	Control N=882	Treatment N = 781	Difference	P-value
	(1)	(2)	(1) - (2)	
Full sample				
Mean ($\bar{x}$)	0.157	0.202	-0.045	0.399
				$diff \neq 0$
Standard Deviation (SD)	1.121	1.039	0.082	0.014**
				$\frac{sd(1)}{sd(2)} > 1$
Pregnant at baseline				
Mean ($\bar{x}$)	0.198	0.127	0.071	0.338
				$diff \neq 0$
Standard Deviation (SD)	1.151	0.964	0.187	0.0001***
				$\frac{sd(1)}{sd(2)} > 1$
Postpartum at baseline				
Mean ($\bar{x}$)	0.114	0.275	-0.161	0.0353
				$diff \neq 0$
Standard Deviation (SD)	1.089	1.104	-0.015	0.607
				$\frac{sd(1)}{sd(2)} > 1$

Notes: The table reports the distribution of net changes in fertility preferences Δ_i by treatment status. Mean rows present average changes in Δ_i, while Standard Deviation (SD) rows report dispersion of Δ_i within each group. P-values in the Mean rows test for equality of means between the treatment and control groups, while p-values in the SD rows test whether dispersion is greater in the control group. A significant difference in SD indicates lower preference stability among control women. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

7 Mechanisms

The conceptual framework set out in Section 2 identifies three channels through which the intervention may shape preference stability: a deliberation channel that operates through the reproductive life-cycle, an information channel that operates through the marginal value of counseling content, and an access channel that operates through the implementability of preferences. The deliberation channel is supported by the heterogeneity in effects by baseline reproductive status documented in Section 6: stability gains concentrate among women who were pregnant at baseline. In this section, we test the remaining two channels through heterogeneity analyses and examine the implications of preference stability for downstream contraceptive behavior.

7.1 Information and Access Channels

To test the information and access channels, we estimate the same two-stage stability model used in our main analysis on subgroups defined by women's baseline characteristics. Heterogeneity in the treatment effect on within-woman variability across these subgroups speaks to which channels are operating.

The information channel operates through the counseling component of the intervention, which provides individualized information about contraceptive methods, side effects, and birth spacing across up to six visits. The counseling adds more new information for women who started with less prior reproductive knowledge. We use women's highest completed level of education — at most primary versus secondary or higher — as a proxy for prior knowledge, and predict a larger stability effect among the lower-education subgroup. The access channel operates through the intervention's transport and cost-relaxation components, which reduce the practical barriers to obtaining contraception. These components matter most for women who face the largest access constraints to begin with — in particular, those who live farther from a family planning facility. We proxy binding access constraints with the neighborhood mean of self-reported distance to the nearest facility, split at the sample median.[7] The framework predicts a larger stability effect in high-distance neighborhoods. Since the framework treats the deliberation channel as moderating the operation of the other two, we report results pooled across the full sample and separately by baseline reproductive status.

Table 3 reports the results. Pooled across the full sample (Panel A), the treatment effects on stability are uniformly negative but small and statistically indistinguishable from zero

[7] We aggregate distance to the neighborhood level to address substantial individual-level missingness in the self-reported variable.

across all four specifications. Among pregnant women with at most primary education, the intervention reduces preference variability by 22.6 percent (-0.256, $p < 0.01$); among those with secondary or higher education, the estimated effect is essentially zero. The contrast is in the direction predicted by the information channel: the intervention produces stability gains where its counseling content has the greatest marginal informational value, namely among women with less prior reproductive education.

The access-channel test follows the same pattern. In high-distance neighborhoods, where access constraints are more binding, the treatment effect on stability is -0.191 ($p < 0.05$), corresponding to a 17 percent reduction in within-woman variability. In low-distance neighborhoods, the effect is small and statistically indistinguishable from zero. Both information and access channels thus operate in the predicted direction within the pregnant subgroup, where reproductive planning is salient.

The within-postpartum tests (Panel C) provide a useful negative control. Among women postpartum at baseline, the treatment effect on stability is statistically indistinguishable from zero in three of the four subgroups. The exception is postpartum women with at most primary education, who show a marginally significant positive coefficient ($+0.147$, $p < 0.10$) — the opposite direction from the prediction. We treat this estimate cautiously, given that the postpartum subsample shows no overall effect on stability in our main analysis and the high-distance subgroup shows no comparable pattern.

Taken together, the heterogeneity tests support both the information and access channels — but only among women pregnant at baseline, where reproductive planning is most salient. This matches the framework's prediction that reproductive-life-cycle stage moderates the operation of the other two channels: when planning is salient (during pregnancy), the intervention's counseling and access components translate into more stable preferences; when it is not (postpartum), they do not. Within the pregnant subgroup, the effects concentrate among women who have the most to learn from counseling (lower-educated) and face the strongest access constraints (higher-distance neighborhoods).

Table 3: Heterogeneity in Treatment Effects on Preference Stability

	Education level		Distance to health facility	
	(1)	(2)	(3)	(4)
	$\leq$ Primary	$>$ Primary	High distance	Low distance
Panel A: Full sample				
Treatment	-0.052	-0.036	-0.067	-0.020
	(0.064)	(0.067)	(0.061)	(0.074)
N	3161	2408	3642	1927
Panel B: Pregnant at baseline				
Treatment	-0.256***	0.025	-0.191**	-0.059
	(0.093)	(0.093)	(0.088)	(0.100)
N	1655	1186	1893	948
Panel C: Post-partum at baseline				
Treatment	0.147*	-0.107	0.039	0.020
	(0.089)	(0.096)	(0.083)	(0.106)
N	1506	1222	1749	979

Notes: The table reports treatment effects on within-woman preference stability by education and access subgroups. The outcome is $|\hat{\varepsilon}_{it}|$, the absolute value of the residual from a first-stage DiD specification with woman and survey-wave fixed effects (Equation 3); coefficients are from a second-stage Gamma GLM with a log link, estimated separately for each cell. Education level (columns 1–2) is the woman's highest completed level of education, split into two categories: $\leq$ Primary refers to women with at most primary education; $>$ Primary refers to women with secondary or higher education. Distance to health facility (columns 3–4) is the neighborhood-mean of self-reported distance (in km) to the nearest family planning facility, split at the sample median. High distance indicates above-median neighborhoods, where access constraints are more binding; Low distance indicates neighborhoods at or below the median. Distance is measured at the neighborhood level to address individual-level missingness in the self-reported variable. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

7.2 Stability and Contraceptive Use

A finding that an intervention affects preference stability is conceptually interesting on its own, but its empirical significance depends on whether stability matters for behavior. We therefore turn to the most direct behavioral expression of fertility preferences: contraceptive use. Two questions motivate the analysis. First, descriptively, is within-woman preference stability associated with contraceptive behavior in our sample? And second, does the intervention's effect on contraceptive use — previously documented by Karra et al. (2022) — partly operate through preference stability?

To address the first question, we estimate associations between the within-woman instability index and five endline contraceptive outcomes: current use of any modern method, long-acting reversible contraception (LARC), implant use, injectable use, and the probability of a subsequent pregnancy. Each estimate is from an OLS regression at endline, controlling for baseline ideal family size, the baseline level of the outcome, and the same demographic covariates as in our main analysis. The estimates are descriptive: preference stability and contraceptive behavior may be jointly determined by unobserved factors, and we do not interpret them as causal effects of stability on choice. The pattern of associations, however, is informative about whether stability is associated with the behaviors the intervention affects.

Table 4 reports the results. In the full sample (Panel A), women with more variable fertility preferences are significantly less likely to be using any modern contraceptive method at endline: a one-unit increase in the within-woman mean of $|\hat{\varepsilon}_{it}|$ corresponds to a 6.3 percentage-point reduction in current FP use ($p < 0.01$). Injectable use shows a similar, marginally significant negative pattern (-0.053, $p < 0.10$), while LARC, implant use, and the probability of a subsequent pregnancy are essentially unrelated to instability.

This headline result conceals a sharp difference between subgroups. Among women pregnant at baseline (Panel B), the association between instability and current contraceptive use is statistically indistinguishable from zero. Among women postpartum at baseline (Panel C), it is roughly twice as large in magnitude and statistically significant (-0.084, $p < 0.05$). The injectable result follows the same pattern. Two features of the setting help explain the contrast. By construction, women pregnant at baseline were not using contraception at study entry: their endline contraceptive use therefore captures uptake *following the index birth*, conditional on a common starting point of zero use. Postpartum women, by contrast, faced contraceptive decisions throughout the panel period, and their endline use reflects continuation and switching choices for which preferences — and their stability — have meaningfully greater behavioral scope. Where contraceptive choice is more actively in play, the association between stability and choice emerges.

Two further patterns are worth noting. The planning-horizon hypothesis — that women

with more variable preferences would avoid commitment to long-acting methods — finds no support: LARC and implant coefficients are essentially zero across all three panels. The descriptive link between stability and contraceptive behavior is not method-specific; it operates on the broader margin of contraceptive engagement rather than on the choice between long- and short-acting methods. More importantly, the subgroup in which stability most strongly predicts contraceptive behavior (postpartum women) is *not* the subgroup in which the intervention most strongly shifts stability (pregnant women, as documented in Section 7.1). This non-overlap is a substantive feature of the setting in its own right: preference stability matters most for contraceptive choice where choice is most freely determined, but the intervention shifts stability most where pregnancy makes reproductive planning salient.

Table 4: Preference Stability and Contraceptive Use

	(1) Current FP	(2) LARC	(3) Implant	(4) Injectable	(5) Next preg.
Panel A: Full sample					
Within-woman instability	-0.063***	0.002	0.001	-0.053*	0.008
	(0.024)	(0.028)	(0.025)	(0.031)	(0.017)
N	1663	1663	1663	1663	1598
Panel B: Pregnant at baseline					
Within-woman instability	-0.046	0.014	0.012	-0.042	0.007
	(0.034)	(0.041)	(0.037)	(0.043)	(0.018)
N	838	838	838	838	773
Panel C: Post-partum at baseline					
Within-woman instability	-0.084**	0.007	0.003	-0.080*	0.010
	(0.036)	(0.036)	(0.034)	(0.045)	(0.029)
N	825	825	825	825	825

Notes: The table compares the intervention's effects on five endline contraceptive outcomes with and without the inclusion of a within-woman preference-stability mediator. Panel A reproduces the adjusted treatment effects from Karra et al. (2022), regressing each outcome on the treatment indicator at endline with controls for the baseline level of the outcome and demographic covariates. Panel B adds the within-woman instability index — the woman-level mean of $|\hat{\varepsilon}_{it}|$ across waves — as an additional regressor. The change in the treatment coefficient between panels indicates the share of the treatment effect statistically associated with the inclusion of the stability mediator. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

We now turn to the second question. To assess whether the intervention's effects on

contraceptive use operate through preference stability, we follow an informal decomposition approach. For each of the five outcomes considered above, we re-estimate the intent-to-treat treatment effect first with the standard set of baseline controls (replicating the specification of Karra et al. (2022)) and then with the within-woman instability index $|\hat{\varepsilon}_i|$ added as an additional regressor. If the intervention's effect on contraception operates substantially through preference stability, the treatment coefficient should attenuate when the mediator is added. We do not interpret this comparison as identifying a causal mediation effect, given the well-known difficulties in separately identifying variance and level mechanisms when both may respond to the same intervention; we treat it as a descriptive test of whether the two empirical effects are statistically linked.

Table 5 reports the comparison. Panel A reproduces the adjusted treatment effects from Karra et al. (2022): the intervention significantly increases current modern FP use by 5.9 percentage points ($p < 0.01$), LARC use by 5.4 points ($p < 0.05$), and implant use by 4.3 points ($p < 0.05$), while reducing the probability of a subsequent pregnancy by 3.9 points ($p < 0.01$). Panel B adds the within-woman instability index. The treatment coefficients differ between panels by at most 0.001 in absolute terms — well within their standard errors of approximately 0.02 — and the instability index itself shows no statistically significant association with any contraceptive outcome once standard controls are included. The intervention's effects on contraceptive use do not statistically operate through changes in preference stability.

This null decomposition formalizes the non-overlap pattern that emerged in the descriptive analysis. The intervention shifts preference stability (concentrated among women pregnant at baseline) and contraceptive behavior (documented across the sample) through distinct mechanisms operating on different subgroups of women. The pattern is consistent with our conceptual framework: the deliberation channel drives the stability effect — pregnancy makes reproductive planning salient, and counseling can then consolidate fertility goals — while the information and access channels likely drive most of the contraceptive uptake effect documented by Karra et al. (2022). The stability gains we document therefore represent a substantive contribution of the intervention independent of its effects on contraceptive uptake — a separate dimension of reproductive decision-making that the intervention shapes, rather than a step on the path from treatment to contraceptive behavior.

Table 5: Decomposition of Treatment Effects on Contraception via Preference Stability

	(1) Current FP	(2) LARC	(3) Implant	(4) Injectable	(5) Next preg.
Panel A: Replicated treatment effects (Karra et al. 2022)					
Treatment	0.059***	0.054***	0.043**	0.001	-0.039***
	(0.021)	(0.021)	(0.019)	(0.024)	(0.013)
N	1667	1667	1667	1667	1601
Panel B: Adding within-woman preference instability					
Treatment	0.058***	0.054***	0.043**	0.000	-0.039***
	(0.021)	(0.021)	(0.020)	(0.024)	(0.013)
Within-woman instability	-0.040	0.008	0.006	-0.049	0.006
	(0.027)	(0.027)	(0.025)	(0.031)	(0.016)
N	1667	1667	1667	1667	1601

Notes: The table reports descriptive associations between within-woman preference variability and contraceptive outcomes at second-year follow-up. The predictor is the within-woman mean of $|\hat{\varepsilon}_{it}|$ across waves, where $\hat{\varepsilon}_{it}$ are residuals from a first-stage difference-in-differences specification on ideal family size with woman and survey-wave fixed effects (Equation 3). Each column reports an OLS regression of the named outcome on the instability index, controlling for baseline ideal family size, the baseline level of the outcome, and a standard set of baseline covariates (neighborhood fixed effects, age, age at sexual debut, parity, ever use of contraception, education, employment status, religion, and ethnicity). Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

8 Conclusion

This study examines how a comprehensive family planning intervention shapes both the level and the stability of fertility preferences among women in urban Malawi. We document a sharp contrast between aggregate and individual dynamics: the distribution of desired family size is remarkably stable across survey waves, yet nearly 60 percent of women revise their stated preference at least once over the three-year panel, and more than a third revise more than once. This within-woman variability is not reporting noise. It is internally coherent with other fertility measures, tracks the gap between realized and desired fertility, and follows predictable life-cycle trajectories. We therefore treat within-woman variation as a behavioral outcome in its own right. To our knowledge, this is the first experimental study to move beyond mean changes in desired family size and estimate the causal effect

of a FP intervention on the stability of preferences, measuring stability through observed variation across repeated observations rather than self-reported certainty, which is elicited contemporaneously with the preference it is meant to validate and is endogenous to the same contextual factors.

The intervention shapes preferences along both dimensions, with effects that are heterogeneous by baseline reproductive status. It has no average effect on desired family size: modest declines among women pregnant at baseline and increases among postpartum women offset one another in the aggregate. The effect on stability is more pronounced. The intervention significantly reduces within-woman variability among women pregnant at baseline robustly across specifications, across balanced and unbalanced samples, and across a complementary distributional test while leaving postpartum women unaffected. Pregnancy is a period of heightened salience for reproductive decision-making, when the cognitive preconditions for deliberation are already in place, allowing counseling to structure and consolidate a reflective process already underway; postpartum women face competing demands on attention, and forward-looking planning is less salient.

Our mechanism analysis links these effects to the channels in our conceptual framework. Within the pregnant subgroup, stability gains concentrate among women with the least prior reproductive knowledge and the strongest access constraints, consistent with the information and access channels operating where the intervention's counseling and constraint-relaxation components carry the greatest marginal value; the deliberation channel moderates both, translating these components into more stable preferences only when reproductive planning is salient. Notably, the subgroup in which stability most strongly predicts contraceptive behavior postpartum women is not the subgroup in which the intervention most strongly shifts stability, and a decomposition confirms that the intervention's documented effects on contraceptive uptake (Karra et al., 2022) do not operate through preference stability. The stability gains we identify are therefore a distinct contribution of the intervention, operating on a separate margin of reproductive decision-making rather than as a step from treatment to contraceptive use.

Conceptually, these findings support the view that fertility preferences are not fixed primitives but are constructed and revised under uncertainty: an intervention can compress within-woman variability without shifting average desired family size, establishing stability as an empirically distinct dimension of fertility preferences. We interpret reduced variability as preference consolidation, on the assumption that instability primarily reflects uncertainty rather than optimal responsiveness. This assumption is not innocuous. In a setting marked by economic and health volatility and shifting intrahousehold bargaining, some preference flexibility may be adaptive, and reduced variability could in principle reflect lost responsive-

ness rather than improved formation; the two cannot be separately identified from stability effects alone. Our evidence that stability is associated with more consistent contraceptive engagement is consistent with the consolidation reading, but we do not claim that greater stability is unambiguously welfare-improving. External validity is also bounded: the sample comprises married women early in their reproductive lives in a single urban context, and the heterogeneity by reproductive status suggests that effects would differ for women at other life stages or in other settings.

The clearest program implication concerns the timing of components. The intervention stabilized preferences only among women who were exposed during pregnancy, which argues for embedding deliberative FP counseling at antenatal contact points rather than treating the postpartum period as the primary window for structured conversations about future reproductive goals. The program's deliberative and access functions need not, however, run on the same schedule. Because the intervention's effects on contraceptive uptake were larger among postpartum women (Karra et al., 2022) and do not operate through stability, the two functions are best matched to the windows where each is most effective: deliberative counseling front-loaded during pregnancy, when it consolidates fertility goals, and access support transport, cost coverage, and reliable supply sustained into the postpartum period, when continuation and switching decisions are most active. Within the pregnant subgroup, stability gains concentrate among women with the least prior knowledge and the strongest access constraints, so where counseling capacity is limited, prioritizing these women maximizes the return; the role of repeated, individualized contact further suggests that consolidation requires sustained engagement rather than one-time information provision.

These results also reframe what counseling-based programs can be expected to accomplish. The intervention did not shift average desired family size, so it is better understood as helping women form and act on their reproductive goals than as altering those goals a meaningful effect on reproductive decision-making in its own right. Capturing this effect requires incorporating preference stability, measured as within-woman variation across a panel, as an evaluation outcome alongside contraceptive use and mean desired fertility. Behavior-based or single-wave evaluations, and cross-sectional measures of subjective certainty, would miss cases in which a program meaningfully shapes preference formation without changing desired fertility.

These implications should be read with the welfare ambiguity noted above. We do not recommend treating preference stability as an objective to be maximized; rather, the defensible aim is narrower to reduce uncertainty-driven instability by relaxing the informational, deliberative, and access constraints that prevent women from forming clear and implementable reproductive goals, rather than to enforce fixed preferences for their own sake. More broadly,

treating preference stability as an outcome bridges demographic and behavioral perspectives on fertility, moving beyond a narrow focus on either stated intentions or realized behavior, and points toward the formation of preferences themselves as a meaningful object of policy and programs.

References

Ashraf, N., E. Field, and J. Lee (2014). Household Bargaining and Excess Fertility: An Experimental Study in Zambia. *American Economic Review 104*(7), 2210–2237. 1

Bachrach, C. A. and S. P. Morgan (2013). A Cognitive-Social Model of Fertility Intentions. *Population and development review 39*(3), 459–485. 5, 8

Badolato, L., S. R. Hayford, and K. B. Guzzo (2025). Multiple dimensions of uncertainty in fertility goals: recent trends and patterns in the United States. *Genus 81*(1), 14. 3, 9

Barker, R. and I. Buber-Ennser (2024). Uncertainty and flexibility of fertility intentions. *Advances in Life Course Research 61*, 100618. 3, 9

Becker, G. S. (1960). An economic analysis of fertility. *Demographic and Economic Change in Developed Countries*. 1, 5

Bhrolchin, M. N. and . Beaujouan (2019). Do People Have Reproductive Goals? Constructive Preferences and the Discovery of Desired Family Size. In R. Schoen (Ed.), *Analytical Family Demography*, pp. 27–56. Cham: Springer International Publishing. 5

Bongaarts, J. (1990). The Measurement of Wanted Fertility. *Population and Development Review 16*(3), 487. 5

Cleland, J., K. Machiyama, and J. B. Casterline (2020). Fertility preferences and subsequent childbearing in Africa and Asia: A synthesis of evidence from longitudinal studies in 28 populations. *Population Studies 74*(1), 1–21. 3

Coombs, L. C. (1974). The Measurement of Family Size Preferences and Subsequent Fertility. *Demography 11*(4), 587–611. 3, 9

Culpepper, S. A. (2010). Studying Individual Differences in Predictability With Gamma Regression and Nonlinear Multilevel Models. *Multivariate Behavioral Research 45*(1), 153–185. 17

Desai, J. and A. Tarozzi (2011). Microcredit, Family Planning Programs, and Contraceptive Behavior: Evidence From a Field Experiment in Ethiopia. *Demography 48*(2), 749–782. 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. 1, 3

Freedman, R. (1997). Do Family Planning Programs Affect Fertility Preferences? A Literature Review. *Studies in Family Planning 28*(1), 1–13. 1

Hayford, S. R. (2009). The Evolution of Fertility Expectations over the Life Course. *De-*

mography 46(4), 765–783. 3

Hayford, S. R. and V. Agadjanian (2017). Determined to stop? Longitudinal analysis of the desire to have no more children in rural Mozambique. *Population Studies 71*(3), 329–344. 1, 3

Heiland, F., A. Prskawetz, and W. C. Sanderson (2008). Are Individuals Desired Family Sizes Stable? Evidence from West German Panel Data: La taille de famille dsire est-elle stable? Analyse de donnes de panel en Allemagne de lOuest. *European Journal of Population / Revue europenne de Dmographie 24*(2), 129–156. 3

Herrera-Almanza, C. and A. Seitz McCarthy (2025). Strategic responses to disparities in spousal desired fertility: experimental evidence from rural Tanzania. *Journal of Population Economics 38*(4), 85. 1, 3

Howe, L. C. and J. A. Krosnick (2017). Attitude Strength. *Annual Review of Psychology 68*(1), 327–351. 6

ICF, N. S. O. N. M. a. (2017). Malawi Demographic and Health Survey 2015-16. *Zomba, Malawi, and Rockville, Maryland, USA. NSO and ICF*. 4

JohnsonHanks, J. (2007). Natural Intentions: Fertility Decline in the African Demographic and Health Surveys. *American Journal of Sociology 112*(4), 1008–1043. 1

Joshi, S. and T. P. Schultz (2013). Family Planning and Women's and Children's Health: Long-Term Consequences of an Outreach Program in Matlab, Bangladesh. *Demography 50*(1), 149–180. 3

Karra, M. and D. Canning (2020). The Effect of Improved Access to Family Planning on Postpartum Women: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols 9*(8), e16697. 10

Karra, M., D. Maggio, M. Guo, B. Ngwira, and D. Canning (2022). The causal effect of a family planning intervention on womens contraceptive use and birth spacing. *Proceedings of the National Academy of Sciences 119*(22), e2200279119. 3, 8, 11, 26, 27, 28, 30, 31

Kodzi, I. A., J. B. Casterline, and P. Aglobitse (2010). The Time Dynamics of Individual Fertility Preferences Among Rural Ghanaian Women. *Studies in Family Planning 41*(1), 45–54. 1, 3, 4

Liefbroer, A. C. (2009). Changes in Family Size Intentions Across Young Adulthood: A Life-Course Perspective. *European Journal of Population / Revue europenne de Dmographie 25*(4), 363–386. 3

Maggio, D., M. Karra, and D. Canning (2024). Family Planning and Childrens Human Capital: Experimental Evidence From Urban Malawi. *Demography 61*(5), 1667–1698. 11

Morgan, S. P. (1981). Intention and Uncertainty at Later Stages of Childbearing: The United States 1965 and 1970. *Demography 18*(3), 267–285. 3, 9

Mueller, M. W., J. Hamory, J. Johnson-Hanks, and E. Miguel (2022). The illusion of stable fertility preferences. *Population Studies 76*(2), 169–189. 1

Nair, N. K. and L. P. Chow (1980). Fertility Intentions and Behavior: Some Findings from Taiwan. *Studies in Family Planning 11*(7/8), 255. 3

Quesnel-Valle, A. and S. P. Morgan (2003). Missing the Target? Correspondence of Fertility Intentions and Behavior in the U.S. *Population Research and Policy Review 22*(5-6), 497–525. 3

Sennott, C. and S. Yeatman (2012). Stability and Change in Fertility Preferences Among Young Women in Malawi. *International Perspectives on Sexual and Reproductive Health 38*(1), 34–42. 1, 3

Tan, P. C. and N. P. Tey (1994). Do Fertility Intentions Predict Subsequent Behavior? Evidence from Peninsular Malaysia. *Studies in Family Planning 25*(4), 222. 3

Trinitapoli, J. and S. Yeatman (2018). The Flexibility of Fertility Preferences in a Context of Uncertainty. *Population and Development Review 44*(1), 87–116. 4, 6

Yeatman, S., C. Sennott, and S. Culpepper (2013). Young Womens Dynamic Family Size Preferences in the Context of Transitioning Fertility. *Demography 50*(5), 1715–1737. 1, 3, 4, 6, 17

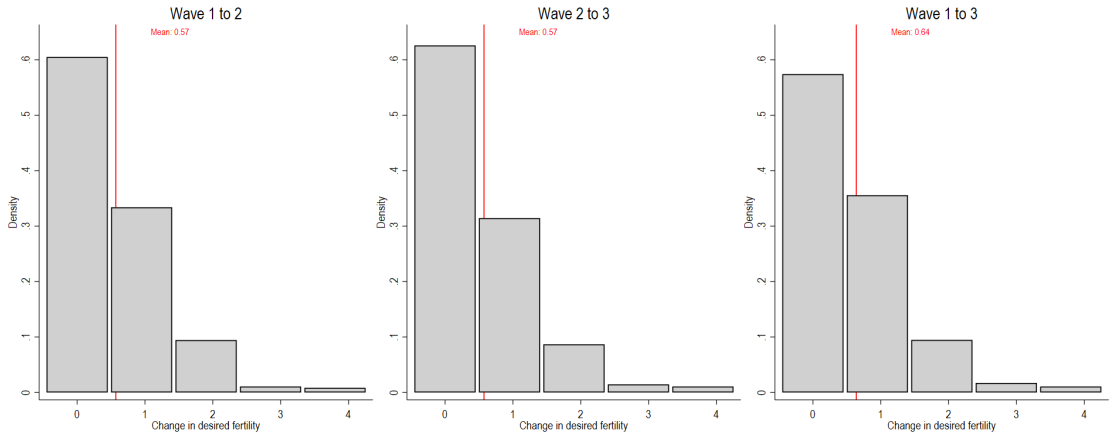
A Appendix

Table A1: Baseline Balance Table

	Full sample	Treatment	Control	Difference
	(1)	(2)	(3)	(3) - (2)
Background characteristics				
Woman's age (years)	24.58	24.66	24.51	-0.148
Primary education (1 = yes)	0.984	0.986	0.982	-0.004
Secondary education (1 = yes)	0.413	0.414	0.412	-0.001
Tertiary education (1 = yes)	0.023	0.019	0.028	0.009
Currently working (1 = yes)	0.096	0.099	0.093	-0.007
Religion	2.909	2.900	2.917	0.018
Ethnicity	1.857	1.875	1.840	-0.035
Age at first sex	18.86	18.90	18.81	-0.094
Currently pregnant (1 = yes)	0.515	0.515	0.516	0.001
Currently using contraception (1 = yes)	0.237	0.239	0.235	-0.003
Ever used contraception (1 = yes)	0.755	0.775	0.736	-0.039**
Number of children	1.738	1.771	1.708	-0.063
Fertility preferences				
Desired number of children	3.187	3.214	3.162	-0.052
Husband same preference (1 = yes)	0.605	0.587	0.622	0.035
Wants another child (1 = yes)	0.572	0.567	0.578	0.011
Preferred waiting time (months)	63.95	63.46	64.39	0.932
Number of observations	2140	1027	1113	2140
Joint p-value				0.260

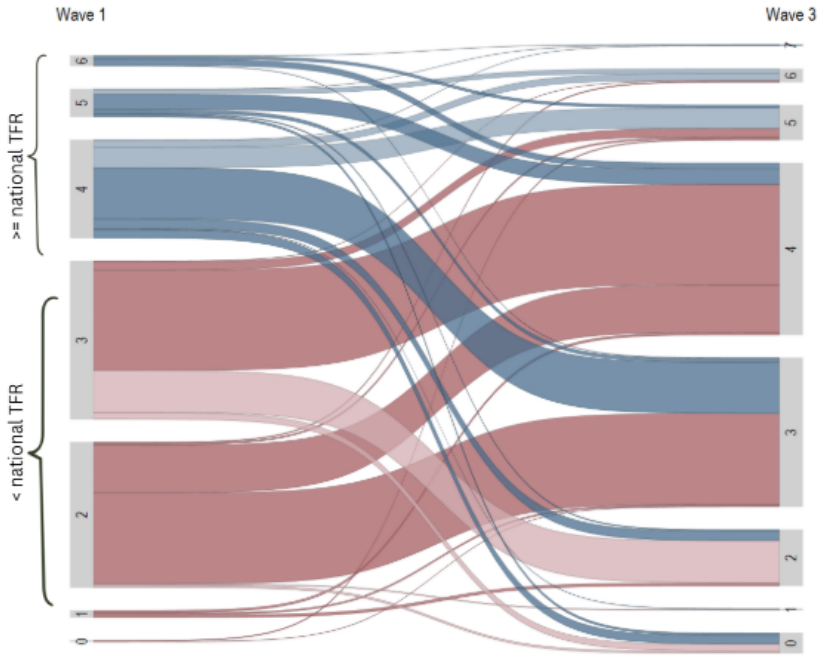
Notes: Means are reported by group, with p-values testing for differences between treatment and control groups. The joint p-value tests for overall differences in baseline characteristics between groups.

Figure A1: Absolute Change in Fertility Preferences across Waves



Notes: The figure displays the distribution of absolute changes in womens reported desired number of children between consecutive survey waves. Absolute changes capture the size of stated preference revisions, regardless of direction, with zero indicating no change and larger values indicating larger adjustments.

Figure A2: Net Revision Trajectories of Fertility Preferences



Notes: The Sankey diagram presents within-woman revisions in desired family size between Wave 1 and Wave 3 among women who revised their fertility preferences at least once during the study period. Nodes represent categories in each wave, and flows connect baseline to endline preference values, with widths proportional to the share of women making each transition. Colors and shading differentiate women with baseline preferences below versus at or above the national total fertility rate.

Table A2: Change Patterns in Fertility Preferences

	Full sample	Age Group			Baseline Parity			
		18-24	25-29	30-35	None	One	Two	≥ 3
Changed at least once	0.594	0.578	0.626	0.599	0.549	0.595	0.597	0.620
Times changed								
Once	0.615	0.638	0.609	0.563	0.722	0.625	0.601	0.558
Twice	0.385	0.362	0.391	0.437	0.278	0.375	0.399	0.442
Direction change								
Increased	0.660	0.697	0.627	0.604	0.641	0.700	0.701	0.585
Decreased	0.340	0.303	0.373	0.396	0.359	0.300	0.299	0.415
Timing								
W1 only	0.322	0.338	0.310	0.296	0.368	0.319	0.282	0.336
W1, reverted back	0.233	0.238	0.226	0.231	0.215	0.247	0.218	0.241
W2 only	0.293	0.300	0.299	0.266	0.354	0.306	0.319	0.223
W1 and W2	0.152	0.124	0.165	0.206	0.062	0.128	0.181	0.201

Notes: The table presents patterns of change in fertility preferences for the full sample and by age group and baseline parity. Percentages are conditional on having revised preferences at least once during the study period.

Table A3: Effects on Fertility Preference Levels, Balanced Sample

	Full sample			Pregnant (baseline)			Postpartum (baseline)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.062	0.056		-0.037	-0.036		0.155**	0.147**	
	(0.048)	(0.047)		(0.068)	(0.066)		(0.068)	(0.067)	
Treatment × Post			0.028			-0.072			0.127*
			(0.044)			(0.058)			(0.066)
Baseline Controls		√			√			√	
Neighborhood FE		√			√			√	
Woman FE			√			√			√
Year FE			√			√			√
Control mean	3.32	3.32	3.32	3.32	3.32	3.32	3.33	3.33	3.33
Number of observations	1664	1662	5276	839	838	2680	825	824	2596

Notes: The table reports the effects of the FP intervention on fertility preference levels for the full sample and by baseline reproductive status. These estimates replicate those presented in **??**, restricted to the balanced sample of women observed in all three data waves.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effects on Fertility Preference Stability: Residual Variance Approach

	OLS		Gamma GLM	
	(1) Unbalanced	(2) Balanced	(3) Unbalanced	(4) Balanced
Panel A: Full sample				
Treatment	-0.019	-0.019	-0.051	-0.049
	(0.017)	(0.020)	(0.048)	(0.049)
Number of observations	5569	4770	5569	4770
Panel B: Pregnant (baseline)				
Treatment	-0.053**	-0.051*	-0.152**	-0.138*
	(0.025)	(0.028)	(0.069)	(0.076)
Number of observations	2841	2388	2841	2388
Panel C: Postpartum (baseline)				
Treatment	0.013	0.007	0.034	0.017
	(0.025)	(0.026)	(0.069)	(0.066)
Number of observations	2728	2382	2728	2382

Notes: The table reports estimates of the effects of the FP intervention on fertility preference stability using the residual-based measure from Equation 4. Columns (1)(2) present OLS estimates, while Columns (3)(4) present Gamma GLM estimates with a log link. Results are presented for balanced and unbalanced samples, for the full sample, and by baseline reproductive status. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: SD Test of Fertility Preference Stability, Balanced Sample

	Control N=838	Treatment N=752	Difference	P-value
	(1)	(2)	(1) - (2)	
Full sample				
Mean ($\bar{x}$)	0.150	0.205	-0.054	0.313
				$diff \neq 0$
Standard Deviation (SD)	1.107	1.034	0.073	0.028**
				$\frac{sd(1)}{sd(2)} > 1$
Pregnant at baseline				
Mean ($\bar{x}$)	0.199	0.130	0.069	0.355
				$diff \neq 0$
Standard Deviation (SD)	1.118	0.964	0.154	0.002***
				$\frac{sd(1)}{sd(2)} > 1$
Postpartum at baseline				
Mean ($\bar{x}$)	0.100	0.277	-0.177	0.023
				$diff \neq 0$
Standard Deviation (SD)	1.094	1.093	0.001	0.494
				$\frac{sd(1)}{sd(2)} > 1$

Notes: The table presents the distribution of net changes in fertility preferences, Δ_i, by treatment status for the balanced sample. Mean rows present average changes in Δ_i, while Standard Deviation (SD) rows report dispersion of Δ_i within each group. P-values in the Mean rows test for equality of means between the treatment and control groups, while p-values in the SD rows test whether dispersion is greater in the control group. A significant difference in SD indicates lower preference stability among control women. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A6: Tests for Selective Attrition: Baseline Characteristics by follow-up status

	Followed up	Attrited	Diff.	SE
	(1)	(2)	(3)	(4)
Treatment assignment (1 = yes)	0.473	0.500	-0.027	(0.025)
Fertility preference	3.18	3.21	-0.028	(0.049)
Other characteristics				
Woman's age (years)	24.90	23.65	1.249***	(0.227)
Currently pregnant (1 = yes)	0.50	0.56	-0.058*	(0.025)
Primary education (1 = yes)	0.99	0.98	0.008	(0.006)
Secondary education (1 = yes)	0.45	0.30	0.150***	(0.024)
Tertiary education (1 = yes)	0.03	0.01	0.017*	(0.007)
Working (1 = yes)	0.10	0.08	0.026	(0.015)
Religion	2.87	3.03	-0.169*	(0.074)
Ethnicity	1.86	1.85	0.004	(0.065)
Age at first sex	18.95	18.57	0.379**	(0.138)
Currently using contraception (1 = yes)	0.24	0.22	0.028	(0.021)
Ever used contraception (1 = yes)	0.77	0.71	0.066**	(0.021)
Number of children	1.81	1.54	0.262***	(0.067)
Wants another child (1 = yes)	0.55	0.63	-0.079**	(0.025)
Preferred waiting time (months)	63.69	64.63	-0.947	(1.503)
Observations	2140			

*Notes:*The table tests for selective attrition along the two dimensions most relevant to our analysis — treatment assignment and baseline fertility preferences — as well as on other baseline characteristics. Columns (1) and (2) report mean values within each group, where *Attrited* refers to women not observed in all panel waves. Column (3) reports the difference (Followed up − Attrited) from a two-sample t-test, with robust standard errors in column (4). Sample: women interviewed at the 2016 baseline wave. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.