BETA
Bureau
d'économie
théorique
et appliquée

Documents de travail

« Norms Behind Closed Doors: A Field Experiment on Gender Norm Misperceptions and Maternal Employment Decisions in Couples »

Auteurs

Marie Boltz, Monserrat Bustelo, Ana María Díaz, Agustina Suaya

Document de Travail n° 2025 – 37

Septembre 2025

Bureau d'Économie
Théorique et Appliquée
BETA

https://www.beta-economics.fr/

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Norms Behind Closed Doors: A Field Experiment on Gender Norm Misperceptions and Maternal Employment Decisions in Couples

Marie Boltz, [*] Monserrat Bustelo [†] Ana María Díaz [‡]
Agustina Suaya [§]

September 15, 2025

Abstract

We study whether pluralistic ignorance about societal and spousal support for maternal employment sustains gender gaps in labor outcomes. We first elicit second-order beliefs from 1,732 cohabiting couples with young children in Bogotá. Personal support for working mothers is almost universal, yet both men and women substantially underestimate others' support, particularly that of men. We then implement a randomized controlled trial delivering personalized information on prevailing attitudes toward maternal employment. The intervention narrowed belief gaps —raising women's estimates of peer support and men's perceptions of their partners' views— while leaving first-order attitudes unchanged. Treated men were 7–8 percentage points (16 percent) more likely than men in the control group to nominate their wives for a career-building course rather than take the course for themselves; women, whose baseline demand was already high, showed no further change. Treated women intensified job-search efforts, and treated men expressed stronger preferences for work-family balance. These results reflect short-run adjustments in beliefs and reported behaviors, measured within weeks of the intervention.

JEL Codes: R41, R42, D62

Keywords: Gender norms, Female Employment, Pluralistic ignorance, RCT.

[*]University of Strasbourg, BETA, CNRS; CEREQ. m.boltz@unistra.fr
[†]Gender and Diversity Division, Inter-American Development Bank. Email: monserratb@iadb.org.
[‡]Department of Economics, Pontificia Universidad Javeriana. Email: a.diaze@javeriana.edu.co.
[§]Gender and Diversity Division, Inter-American Development Bank. Email: agustina.suaya@gmail.com. This study was pre-registered in the AEA Social Science Registry under Trial ID AEARCTR-0014648

1 Introduction

Mothers' labor-force participation remains far below fathers' —despite a clear majority of adults saying they support women's employment (Bursztyn et al., 2023; Frisancho et al., 2023). Although some structural and institutional barriers persist, growing evidence points to social and cultural norms —especially those assigning primary caregiving responsibilities to women— as the main drivers of the enduring gender gap. Recent studies highlight pluralistic ignorance —the widespread misperception of others' beliefs and attitudes— as a key mechanism sustaining these norms (Cameron et al., 2024; Laszlo et al., 2023). When people underestimate their peers' support for working mothers, they conform to a norm they privately reject, perpetuating outdated gender roles. If such mistaken second-order beliefs underpin the participation puzzle, correcting them should realign attitudes with reality and spur behavioral change.

We test whether correcting pluralistic ignorance about community support for working mothers both narrows spouses' misperceptions about societal beliefs and about their partner and improves mothers' labor-market prospects. Concretely, does providing accurate information lead husbands to prioritize their wives for a competitive career-development course and does it increase women's employment aspirations and job-search efforts?

To answer this question, we elicit beliefs among 1,732 cohabiting couples with children under 6 years old, in Bogotá a large, middle-income metropolis. We focus specifically on two types of misperceptions: (i) misperceptions about societal support for maternal employment (community-level beliefs), and (ii) misperceptions about a partner's support (couple-level beliefs).

We designed a three-step randomized controlled trial. First, we combined in-person and phone surveys to elicit (i) personal attitudes and (ii) second-order beliefs about com-

munity and spousal support for maternal employment and intrahousehold task-sharing. Second, we used household-level stratified randomization —stratifying on baseline employment status and beliefs— to assign couples to treatment or control. In treated households, both partners received their own personalized WhatsApp message that contrasted their individual perceptions with accurate neighborhood-level data on support for working mothers; control households received neutral labor-market facts. Because the randomization operates at the couple level while the feedback is individualized, any downstream changes in decision-making provide a clean test of "generalized introspection" within the household (i.e., reflection on one's own social assumptions prompted by accurate community feedback). Third, we recontacted participants two months later to measure belief updating, course allocation, job-search actions, and employment outcomes.

Because engagement with the midline (WhatsApp chatbot) and endline survey was partial, our analysis uses different samples depending on when the outcome was measured. We report which sample is used in each result and correct for selective exposure and attrition using inverse probability weighting (IPW).

Baseline data reveal near-universal support for maternal employment (96% of husbands and 99% of wives), yet both men and women substantially underestimate societal support—by roughly 20 percentage points. These misperceptions are more pronounced among men and extend to intra-household expectations: husbands overestimate their wives' conservatism, while wives underestimate their partners' willingness to share domestic responsibilities. The intervention did not affect first-order attitudes, which were already high, but significantly reduced second-order misperceptions. Treated men and women revised their beliefs about peer and spousal support by 7–9 percentage points, with stronger effects where baseline misalignment was greatest. These belief adjustments were accompanied by meaningful shifts in reported household decisions: treated

men were 16% more likely to prioritize their wives for a competitive career program, and treated women reported increased job search activity. Effects persisted even in households where only one partner received the intervention, underscoring the potential of belief correction within couples to influence labor-related choices.

This study builds on a growing literature documenting that misperceived gender norms depress women's labour-market outcomes. Surveys in sixty countries document gaps of 15–30 percentage points between actual and perceived societal support for working mothers (Bursztyn et al., 2023). Bernhardt et al. (2018) similarly found that men in India significantly overestimate community disapproval of women working, which distorts intra-household bargaining and constrains women's labor market engagement. Bohren et al. (2023) formalise the distinction between genuine tastes and inaccurate beliefs, arguing that policy must target misperceptions, not preferences.

Field experiments confirm that norm-correcting information can shift behaviours. In Saudi Arabia, Bursztyn et al. (2020) showed that informing men of actual support for women's employment increased their willingness to support their wives working, with sustained effects over time. In Indonesia, Cameron et al. (2024) found that women exposed to similar norm-correcting information were 25% more likely to select a career mentoring course over a shopping voucher. Likewise, Laszlo et al. (2023) documented that a norm-shifting intervention in Paraguay led to stronger beliefs in equitable household labor division among female caregivers.

Our study adds three elements to a literature that documents gaps between private support for working mothers and the behaviour that actually prevails. First, we examine pluralistic ignorance in a large Latin American metropolis with relatively high female labor-force participation, illustrating that belief gaps can persist—and shape outcomes—even in contexts where women's economic participation is more established. Second, we randomize and survey *both* partners in 1 732 cohabiting couples, which lets

us track *mutual* misperceptions and intrahousehold decision-making about female LFP —something most single-respondent studies cannot observe. Third, we test whether community-level information alone, delivered through personalized WhatsApp messages, spills over to spousal beliefs and affects individual choices within the couple: whether to keep a career-development course for oneself or offer it to one's partner. By focusing on couples raising young children, our within-household design isolates how updated beliefs translate into maternal employment decisions. This offers insights with potential external validity outside the Latin-American context. Indeed, the arrival of children has been shown to be a critical turning point that sharply widens gender gaps in labor force participation and earnings across diverse settings (e.g., Goldin, 2021; Kleven et al., 2019), making our focus on couples with young children particularly relevant.

The remainder of the paper is organized as follows. Section 2 presents the experimental design and the theory of change. Section 3 describes the data with a focus on the baseline beliefs elicitation. Section 4 outlines the empirical approach, including data collection, randomization, and outcome measures. Section 5 discusses the main results, including belief updating, intra-household dynamics, and job market behaviors. Finally, Section 6 concludes.

2 Experimental Design

Our experimental design involves three sequential stages implemented with a consistent sample of cohabiting couples: (i) a baseline survey to elicit first- and second-order beliefs; (ii) an informational intervention delivered via WhatsApp; and (iii) a follow-up survey to measure belief updating, job search, and labor market aspirations.

Baseline survey. Between July and September 2024 we interviewed 1,732 co-habiting heterosexual couples in Bogotá, i.e., 3,464 adults, with at least one child under six. Data collection was conducted through a combination of in-person and telephone interviews, depending on participants' availability.

The questionnaire included eight statements on gender norms and maternal employment; here we focus on the item "Mothers of children under six should be free to work for pay outside the home." For every statement we collected three measures: (i) first-order belief —the respondent's own stance (agree/disagree)—; (ii) community second-order belief —respondents estimated, on a 0–100 scale, how many out of 100 fathers (respectively, mothers) with young children from Bogotá would agree with the statement—; and (iii) spousal second-order belief —a direct prediction of the partner's stance, asked as "Do you think your partner would agree or disagree with this statement?"

The initial sample was large to be representative of the various economic strata of the population of young parents in Bogotá. This feature was particularly important to be able to calculate the *actual* beliefs of men and women with young children and compare them to the second-order beliefs. We measure pluralistic ignorance as the gap between a respondent's community forecast and the *actual* share of agreement among our representative sample of Bogotà.

The survey also captured household demographics, employment status, and other socioeconomic covariates used to define the randomization strata and to improve precision in the analysis.

Midline survey – Information intervention The intervention phase began on October 15, 2024, and included the entire baseline sample. We randomly assigned the 1,732 households to treatment or control with a stratified design that used two baseline vari-

ables —the wife's employment status and the husband's first-order belief about maternal employment— to keep those predictors balanced across groups.

Participants in the treatment group (866 households) received targeted informational messages via a WhatsApp chatbot designed to target one norm: "Mothers of children under six should be free to work for pay outside the home."

The intervention used a discrepancy-based feedback strategy. Each respondent (i) saw the answer they had given in the baseline survey; (ii) guessed whether that answer matched the population; (iii) received the true share from our own baseline data (presented with an emoji counter); and (iv) rated the gap as "interesting," "irrelevant," or "disappointing." We then repeated the same four-step cycle for men's attitudes, randomizing whether the male or female statistics appeared first.

Control households (866 couples) received placebo content through the same channel, on the same schedule and with the same format, but the content concerned an unrelated topic: attitudes toward corporate subsidies for green transport. Because delivery, frequency, and user interface were identical, any outcome differences isolate the effect of correcting misperceived gender norms.

After the information module, the chatbot invited all participants to claim an online career-development and skill-building course. Crucially, participants were also asked whether they wished to enroll themselves or offer the course opportunity to their spouse or partner, given that only one slot was available per household. This introduced a household-level trade-off into the experimental design.

Engagement with the WhatsApp module was moderate: 1,019 of 3,464 adults (29 percent) clicked through the content: meaning saw the information of the treatment or control groups. In the analysis section we address this selective take-up with inverse-probability weights.

Endline survey Between November and January 2025 we reached 1,388 of the original 3,464 participants by phone —accounting for about 40 percent of the baseline sample. Half of this endline sample had previously participated in the midline (WhatsApp chatbot) survey, while the other half was reached only at endline (See Figure 1).

All endline respondents first completed a module measuring first-order beliefs. Then, because only half of the assigned treatment group had engaged with the WhatsApp chatbot, we re-administered the treatment after the first-order beliefs elicitation to everyone assigned to the treatment group. Interviewers followed a verbatim script that replicated the WhatsApp sequence: they reminded the respondent of his or her baseline guess, presented the true figure for Bogotá, and asked the respondent to classify the gap as "interesting," "irrelevant," or "disappointing." Interviewers repeated the statistic until the respondent chose a label, ensuring comprehension. As a result, all treated participant received the information during the phone call, but treatment intensity varied: some respondents were exposed once (phone only), others twice (WhatsApp and phone). We therefore estimate treatment effects using outcome-specific samples and apply inverse probability weighting to correct for selection into exposure and survey participation. Attrition between baseline and follow-up was substantial (60%) and systematically related to socioeconomic and attitudinal characteristics. This heterogeneity is central to our identification strategy and is addressed consistently in the empirical analysis.

Theory of change. Our analysis tests whether information triggers belief updating, which in turn revises intra-household expectations and ultimately shapes women's labor-market choices. We expect that (i) underestimating community support suppresses both one's own actions and a partner's support; (ii) accurate statistics realign second-order beliefs with reality; and (iii) updated beliefs translate into greater job search, stronger openness to paid work, and increased support for maternal employment. Because part-

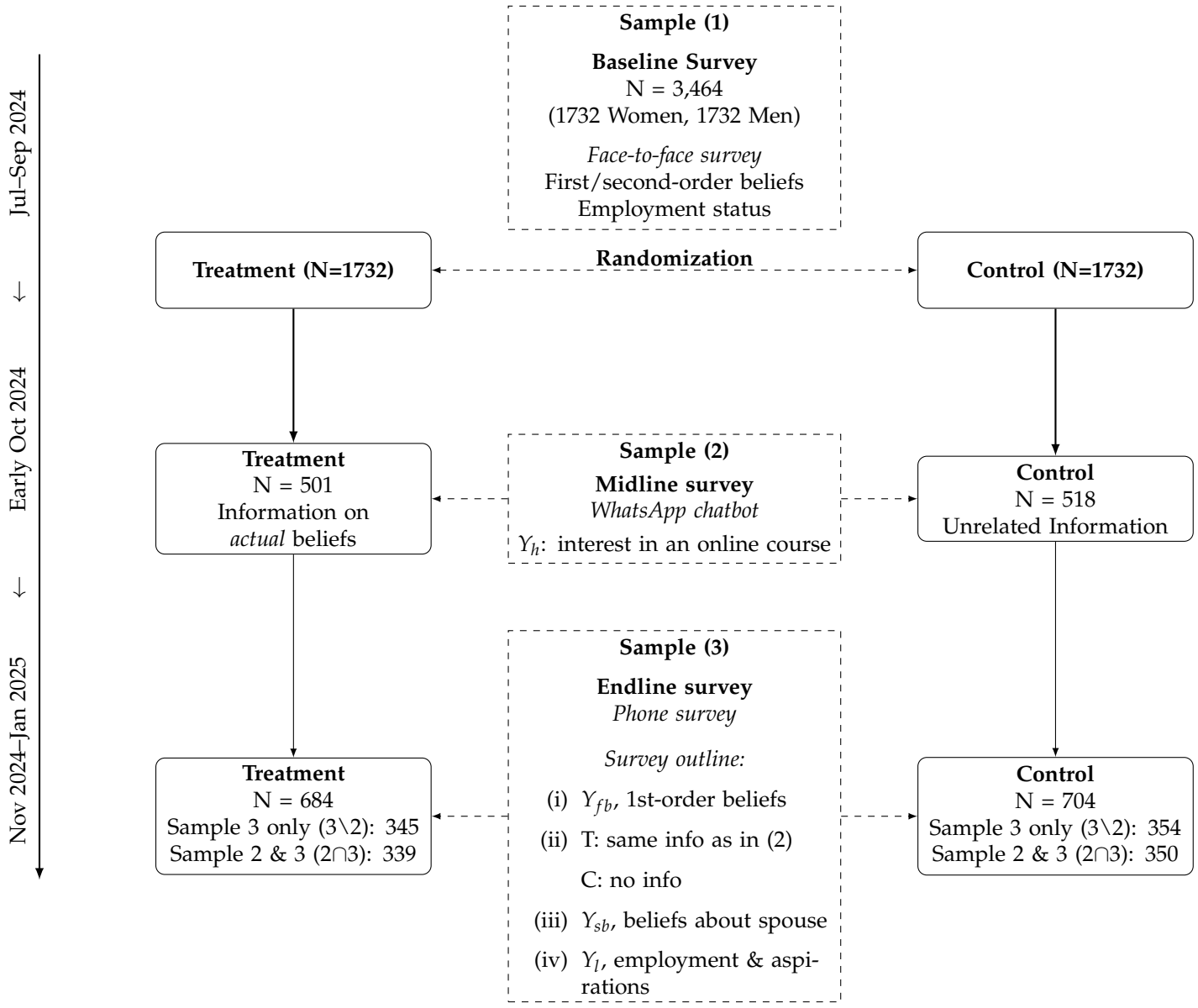
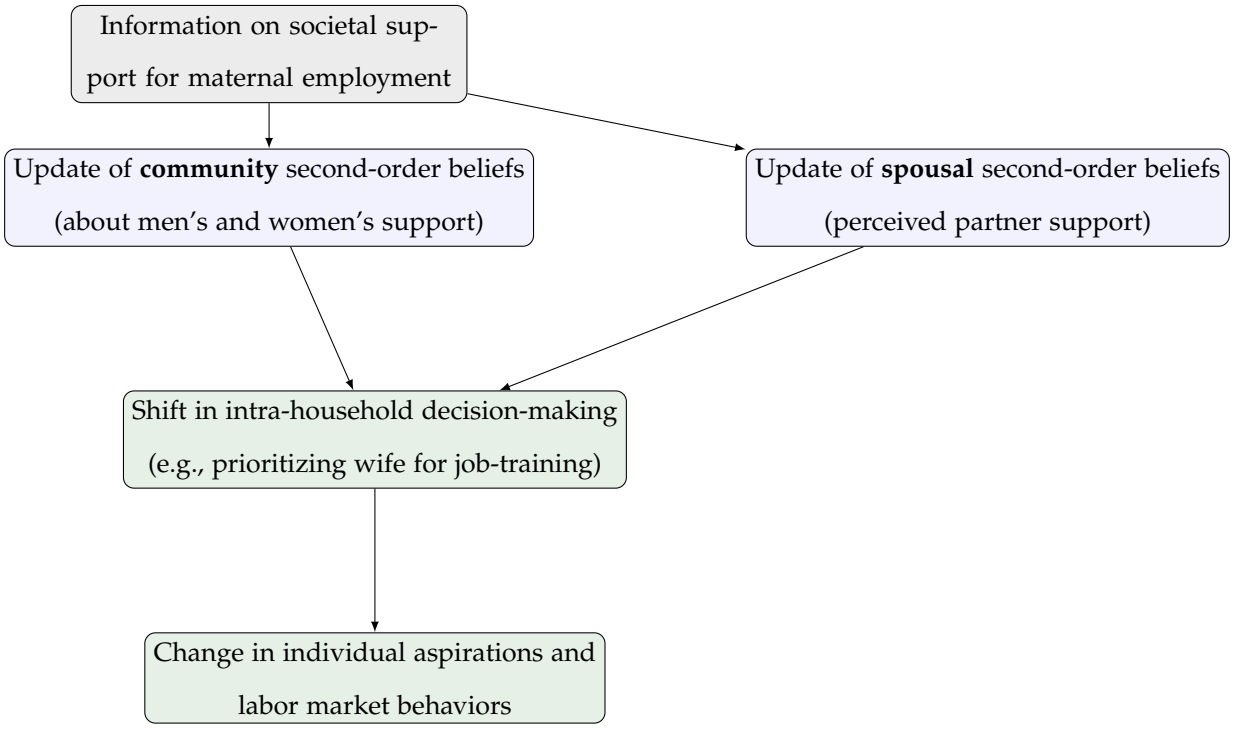


Figure 1. Experimental design and samples

ners often misread each other's views (Boltz et al., 2025), closing these spousal misperceptions may be critical for meaningful change.

The underlying logic of the design hinges on how information affects belief formation and household dynamics. Figure 2 illustrates the mechanism through which the intervention is expected to influence intra-household outcomes. By confronting individuals with accurate data on societal support for maternal employment, the treatment can prompt belief updating along two dimensions: second-order beliefs about community norms, and spousal second-order beliefs—what respondents believe their partner thinks. These revised perceptions may alter intra-household expectations, particularly regarding career prioritization or caregiving support. In turn, shifts in perceived endorsement may affect decision-making, such as nominating the wife for a job-training opportunity, and ultimately influence individual labor market aspirations. This pathway is consistent with *generalized introspection* —that is, belief revision triggered by credible social information, which may extend to adjacent domains such as perceived partner support.

Figure 2. Theory of change: how correcting misperceived gender norms may affect intra-household and labor market outcomes



Notes: The intervention provides respondents with accurate information on prevailing social norms regarding maternal employment. This may trigger belief updating about societal and spousal support, potentially shifting intra-household decisions and individual labor market aspirations.

3 Data description

The baseline covers 1,732 co-habiting heterosexual couples in Bogotá—3,464 adults, each household with at least one child younger than six. Table 1 presents descriptive statistics on household-level attributes. Two-thirds of interviews took place by phone, the remainder in person. The average household has 3.8 members, including 1.1 children under 6 years old; 28 percent report an unenrolled child under six and 32 percent have a member who needs permanent care. In terms of income, 28% of households fall into the low-income category, 60% are middle-income, and 12% are high-income, mirroring Bogotá's income distribution.

Table 1. Household and Care Context

	Mean	Std. Dev.
Survey modality		
In-person interview	0.320	0.47
Telephone interview	0.680	0.47
Household composition		
Household size	3.810	1.05
Children under 6	1.129	0.37
Any child <6 not enrolled	0.276	0.45
Member needs permanent care	0.320	0.47
Income category		
Low income	0.280	0.45
Middle income	0.600	0.49
High income	0.120	0.33
Observations	1,732	

Notes: Means are proportions unless stated otherwise. "Any child <6 not enrolled" equals 1 if the household reports at least one unenrolled child under six. "Member needs permanent care" equals 1 if any household member requires ongoing daily assistance.

Finally, Table 2 presents individual-level characteristics by gender. Male partners are, on average, 2.8 years older than their female counterparts and more likely to have low educational attainment. Employment rates are significantly higher for men (90.5%) compared to women (52%), and men work an average of 11 more hours per week. However, women report greater job flexibility and higher compatibility between work and family life (83.6% vs. 77.4%). Women are also more likely to be inactive (41.7% vs. 4.5%) and actively seeking work (16.2% vs. 10.6%). Men are slightly more likely to be preparing to start a business and to report satisfaction with their current employment situation (31% vs. 24.9%).

Table 2. Comparison of Characteristics by Gender

	Husband		Wife		Difference
Variable	Mean (SE)	N	Mean (SE)	N	Diff
Demographics					
Age	34.859 (0.21)	1732	32.033 (0.18)	1732	2.826***
Education Levels					
Low Education	0.143 (0.01)	1732	0.108 (0.01)	1732	0.035***
Medium Education	0.695 (0.01)	1732	0.711 (0.01)	1732	-0.016
High Education	0.162 (0.01)	1732	0.181 (0.01)	1732	-0.020
Employment Status					
Employed	0.905 (0.01)	1732	0.520 (0.01)	1732	0.385***
Unemployed	0.050 (0.01)	1732	0.063 (0.01)	1732	-0.013
Inactive	0.045 (0.01)	1732	0.417 (0.01)	1732	-0.372***
Working Hours	48.730 (0.36)	1666	37.639 (0.49)	1277	11.091***
Job Flexibility					
High Flexibility	0.236 (0.01)	1666	0.330 (0.01)	1277	-0.094***
Some Flexibility	0.272 (0.01)	1666	0.315 (0.01)	1277	-0.043**
No Flexibility	0.489 (0.01)	1666	0.351 (0.01)	1277	0.138***
Work-Family Compatibility					
Compatibility Score	0.774 (0.01)	1666	0.836 (0.01)	1277	-0.062***
Job Search Activities					
Looking for a job	0.106 (0.01)	1709	0.162 (0.01)	1706	-0.055***
Preparing to start a business	0.091 (0.01)	1709	0.072 (0.01)	1706	0.019**
No, but I would like to	0.493 (0.01)	1709	0.518 (0.01)	1706	-0.025
No, satisfied	0.310 (0.01)	1709	0.249 (0.01)	1706	0.062***

Notes: Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Education levels: *Low* — None, Preschool, Primary (Grades 1–5), and Lower Secondary (Grades 6–9); *Medium* — Academic High School (Grades 10–11), Technical High School, Normal School, Professional Technician, and Technologist; *High* — University, Specialization, Master's Degree, or Doctorate. Sample sizes vary by item due to non-response or applicability (e.g., only employed respondents were asked about working hours and job flexibility).

3.1 Baseline beliefs

Table 3 presents beliefs regarding whether mothers of children under six should have the freedom to work outside the home. Support for this norm is high—88.5% among husbands and 90.5% among wives. However, both husbands and wives significantly underestimate male societal support, with husbands estimating 61.0% and wives 55.7%. Perceptions of women's support remain high and aligned across genders (80%).

To assess within-household misperception, we compare each respondent's belief about their spouse with the spouse's own reported belief. The resulting within-couple gap is roughly four percentage points for men and two for women. This yields a smaller but still meaningful discrepancy: on average, individuals overestimate their partner's support by about 2 percentage points. This suggests that informational frictions may persist even within couples, potentially shaping labor market decisions in subtle but consequential ways.

The placebo norm, by contrast, shows near-universal agreement and a negligible gender gap, confirming that misperceptions concentrate on gender-role content, not on neutral topics. These gaps—large at the societal level and still present within couples—define the informational constraint the intervention seeks to relax.

Table 3. Baseline beliefs for the target and placebo norms

	Husbands (% agreeing) (1)	Wives (% agreeing) (2)	Diff. (1-2)
A. Mothers of children <6 should be free to work			
First-order (own view)	88.50 (0.76)	90.50 (0.68)	–2.00**
Second-order: share of *men* who agree	60.98 (1.19)	55.70 (1.30)	5.28***
Second-order: share of *women* who agree	79.61 (1.10)	80.01 (1.10)	–0.40
Spousal second-order	93.90 (0.52)	89.90 (0.70)	4.10***
B. Placebo norm: companies should subsidise public transport			
First-order (own view)	93.50 (0.60)	94.90 (0.50)	–1.40***

Notes: Weighted means; robust standard errors in parentheses. "Diff." reports two-sided t-tests for equality of male and female means. The placebo norm was asked pre-randomization but shown again only to the control arm during the chatbot. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

We also explored other gender norms and find that progressive private attitudes coexist with large, systematic misperceptions —especially about men—at both the community and couple levels. Table 4 Panel A shows that nearly all respondents endorse gender-equitable statements: 96 % of men and 99 % of women say "women should be allowed to work," and support is only eight to ten points lower when the statement

names mothers of under-six children (88 % and 91 %). Endorsement of equal caregiving is virtually universal. Yet a sizable minority remains attached to traditional norms: one in three think "children suffer if the mother works," and one in four accept male job priority or fear that a higher-earning wife will cause problems.

Panel B present societal beliefs; the rows labeled "other men" show pervasive pluralistic ignorance. Husbands think barely 63 % of fathers support women's work; wives place the figure below 58 %. The true figure (Panel A) is 88–96 %, so respondents underestimate men's support by roughly 30–40 percentage points. Misperceptions are even larger for equal caregiving (50 % vs. 98 %). Re-asking the same questions under the *social-desirability-bias* (SDB) prompt—"if other men answered honestly"—barely moves the averages (< 1 pp), confirming the gaps reflect information, not strategic reporting. Perceived support among women is closer to reality but still understated by 10 pp on average. Finally, Panel C shows spousal beliefs. Inside the household, wives underestimate husbands' support for equal caregiving by 10 pp (88 % vs. 98 %), while husbands slightly over-estimate wives' support for mothers working. For the income-gap norm, both sexes think the partner is 2–3 pp more conservative than reported.

Table 4. Baseline beliefs by belief type and respondent gender

Statement	Husbands Mean (s.e.)	Wives Mean (s.e.)	Diff. (H–W)
A. First-order beliefs (own view)			
Women should be allowed to work	96.1 (0.32)	98.6 (0.18)	–2.5***
Fathers should share caregiving equally	97.9 (0.31)	98.9 (0.22)	–1.0*
Children suffer if mother works	28.3 (1.09)	34.0 (1.14)	–5.7***
Men should have job priority in scarcity	28.2 (1.09)	23.9 (1.04)	4.3***
Problems if wife earns more than husband	22.1 (1.01)	26.4 (1.08)	–4.3***
B1. Second-order beliefs: *other men*			
Women should work	62.7 (1.17)	57.5 (1.28)	5.2***
Equal caregiving	61.9 (1.26)	49.9 (1.34)	12.0***
Higher-earning wife causes problems	57.2 (1.27)	59.8 (1.28)	–2.6**
B2. Second-order beliefs: *other men* (SDB framing)			
Women should work	62.2 (1.17)	58.0 (1.28)	4.2***
Equal caregiving	60.6 (1.26)	50.2 (1.33)	10.4***
Higher-earning wife causes problems	56.0 (1.28)	58.7 (1.29)	–2.7**
B3. Second-order beliefs: *other women*			
Women should work	80.3 (1.00)	80.5 (1.00)	–0.2
Equal caregiving	86.6 (0.88)	88.2 (0.84)	–1.6**
Higher-earning wife causes problems	55.5 (1.27)	58.6 (1.25)	–3.1***
C. Spousal second-order beliefs			
Partner – Equal caregiving	98.5 (0.24)	88.3 (0.79)	10.2***
Partner – Higher- earning wife causes problems	24.8 (1.02)	27.6 (1.07)	–2.8**

Notes: Weighted means; robust standard errors in parentheses. "SDB framing" repeats the second-order item with the prompt "if other men answered honestly" to gauge social desirability bias (SDB). Items "children suffer" and "male job priority" were not asked under SDB or spousal frames; cells are omitted. "Diff." gives the two–sided t-test for equality of male and female means. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. True population values are computed from respondents' first-order beliefs within the same sample, separately for men and women.

4 Empirical strategy

We estimate *treatment–on–the–treated* effects. The exposure indicator equals one if the respondent was in the treated group and it equals zero otherwise.

We analyze four groups of outcomes that follow the causal pathway outlined in Figure 2. Because each outcome was measured at a different stage of the study—and not all individuals completed both the midline and endline surveys—the final sample varies by outcome and is defined by treatment exposure and survey participation. These samples

are described below and visually summarized in Figure 1.

- **Private beliefs** and **community second-order beliefs** (Y_b) were collected at the start of the endline survey, before the treatment reinforcement. The relevant sample includes respondents who participated in both the midline WhatsApp survey (Sample 2) and the endline phone survey (Sample 3). Belief updating focuses on two key perceptions: (i) support for maternal employment among mothers and fathers in Bogotá, and (ii) support for equal caregiving.

- **Spousal perceptions** (Y_s), which capture the accuracy of beliefs about the partner's views, were elicited after the treatment reinforcement during the endline. These outcomes are analyzed on the full endline sample (Sample 3) and the overlap sample (2) $\cap$(3).

- **Intra-household allocation** (Y_h) refers to the respondent's choice of whether to claim or transfer the career-development course opportunity. This decision was recorded at midline and is analyzed using Sample 2.

- **Labor market behaviors** (Y_l) —including job search effort, labor market aspirations and work and family balance— were measured at endline, two to four months after treatment at midline and reflect short-run responses to the intervention. The relevant sample includes respondents who completed both the midline and endline surveys (Sample 2 $\cap$ 3).

For every outcome we fit the linear model

$$y_i = \beta_0 + \beta_1 \,\text{Treatment}_i + \beta_2 \,\text{Female}_i + \beta_3 \left(\text{Treatment}_i \times \text{Female}_i\right) + \mathbf{X}_i'\gamma + \varepsilon_i, \quad (1)$$

where $\mathbf{X}_i$ contains the four stratification dummies. All regressions include statra fixed effects, cluster standard errors at the household level, and we report exact permutation

Fisher p-values computed within strata. Men are the reference group, so β_1 measures the average exposure effect for men and $\beta_1+\beta_3$ the effect for women.

Evidence of Selective Exposure and Attrition Table 5 regresses the exposure to treatment indicator on baseline covariates for each sample. Exposure is orthogonal to nearly all observables, with only small imbalances—household size, respondent age, and a few belief variables—appearing in specific subsamples.

Table 5. Balance Test for Different Samples (Dependent Variable: Treated = 1)

Variable	**Midline Survey** Sample (2)	**Endline Survey** Sample (3)	**Both Surveys** Sample (2)∩(3)	**Endline Only (No Midline)** Sample (3)\(2)
Household members		0.126*		0.183*
	(0.071)		(0.105)	
Age		0.012*		
		(0.007)		
Second-order beliefs (women):			0.008*	
Women work			(0.005)	
First-order belief:		–0.386*		
Moms under 6 work		(0.199)		
Second-order beliefs (women):				–0.009*
Moms under 6 work				(0.005)
Second-order beliefs (partner):	–0.432*	–0.391*	–0.587**	
Moms under 6 work	(0.232)	(0.216)	(0.298)	
Observations	1,019	1,388	689	699

Notes: Each column reports coefficients from a separate OLS regression in which the dependent variable equals 1 if the individual was treated. All regressions include the full set of stratification controls (omitted for brevity). This table shows only covariates that are statistically different from zero in at least one specification ($p < 0.1$). Standard errors are clustered at the household level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Panel attrition was substantial: only 40% of original participants completed the endline survey. Table 12 (Appendix) documents that retention is systematically correlated with baseline covariates. Respondents who remain are disproportionately female, marginally older, and drawn from larger households that report a permanent-care obligation. The composition of the panel also shifts up the socioeconomic gradient: coefficients on urban strata 1–4 are uniformly negative, indicating that the post-attrition sam-

ple over-represents relatively affluent neighbourhoods. Finally, surviving respondents display more progressive priors on women's labour-force participation, as reflected in the positive association with first order beliefs of women working and the negative association with second order beliefs of mothers working. This underscores the importance of adjusting for selection in all analyses.

These patterns suggest that the final sample over-represents individuals with stronger labor market attachment and more egalitarian views, which could bias estimates downward if these groups are less responsive to belief updating.

To address potential selection into treatment exposure and differential attrition, we implement inverse probability weighting (IPW) as a robustness strategy. We estimate each respondent's probability of receiving the treatment and remaining in the final sample using a single-step propensity model. This model includes the set of baseline covariates found to be significantly associated with either exposure or retention—specifically, individual demographics, household composition, and first- and second-order beliefs about maternal employment elicited at baseline. Hence, in all following tables, we report IPW results alongside unweighted OLS estimates that control for key unbalanced covariates.

Treated individuals are weighted by the inverse of their predicted probability of being treated and retained; untreated individuals are weighted by the inverse of one minus that probability. All IPW estimates are computed separately by gender and are reported alongside unweighted OLS results. Standard errors for the IPW results are computed via nonparametric bootstrap.

Table 6 provides a summary mapping each outcome category to its corresponding estimation sample, clarifying how survey timing and treatment exposure define the analytic samples.

Table 6. Mapping of Outcome Variables to Estimation Samples

Outcome Type	Outcome Variables	Estimation Sample
Beliefs about Society	First-order and second-order beliefs about societal support for maternal employment and equal task sharing	Sample 2 ∩ 3: Respondents who engaged with the WhatsApp chatbot and completed the endline phone survey.
Beliefs about Spouse	Second-order beliefs about partner views and misperceptions about partner support	Sample 2 ∩ 3: Measured after reinforced exposure at the start of the endline survey.
Course Enrollment Preferences	Wife/self prioritization, self-reported interest in attending, and beliefs about partner interest	Sample 2: All respondents who completed the midline WhatsApp module, regardless of later survey completion.
Labor Market Behavior and Attitudes	Changed job / started business Desire to improve labor market status Preference for work–family balance	Three groups analyzed separately: (i) individuals treated during WhatsApp midline and followed up at endline (Sample 2 ∩ 3), (ii) individuals indirectly exposed via a treated partner, and (iii) individuals treated only at endline (Sample 3 \ 2).

Notes: Sample 2 comprises respondents who engaged with the WhatsApp intervention during the midline module. Sample 3 includes individuals who completed the phone-based endline survey. Outcomes are matched to samples depending on timing of measurement and treatment exposure (direct, indirect, or placebo).

5 Results

We present the results to mirror the hypothesized causal pathway described earlier. If exposure to accurate normative information prompts individuals to revise their expec-

tations about societal or spousal support, these revised beliefs may translate into shifts in household decisions and labor market aspirations. Accordingly, we first test whether the intervention altered beliefs —both about Bogota and about one's partner. We then examine whether these belief changes affected within-household decisions, such as the prioritization of the wife for a career-enhancing opportunity. Finally, we assess whether these changes translated into behavioral intentions or actions in the labor market.

Changes in beliefs

First-order beliefs and community second-order beliefs, Y_b. We first explore whether providing the population share that approves mother's paid work has an effect on beliefs. Specifically, we test whether a one-time *WhatsApp*-based informational intervention alters first- and second-order beliefs. These outcomes were measured at the beginning of the follow-up phone survey —prior to any additional information being delivered— ensuring that treated individuals were exposed only once, via the chatbot. The estimation sample includes 689 individuals who completed both the midline WhatsApp survey (Sample 2) and the endline phone follow-up (Sample 3); that is, Sample $2 \cap 3$.

Table 7 presents the results. Column (1) shows treatment effects on respondents' own attitudes (first-order beliefs). Columns (2) and (3) report effects on their estimates of the share of *other* men and women, respectively, who support maternal employment (second-order beliefs). Columns (4) and (5) present the results when the dependent variable is a binary "misperception" indicator that equals one if the respondent still underestimates the true support level and zero otherwise. Panel A shows estimates from OLS regressions based on equation 1 controlling by stratification fixed effects, and baseline variables. Panel B presents inverse probability weighted (IPW) estimates separately by gender, which account for differential engagement and attrition across treatment groups.

We find no evidence that the intervention influenced first-order beliefs (column 1). Participants' own support for maternal employment remained unchanged (column 1). Columns (2) and (3) present treatment effects on second-order beliefs—participants' estimates of the proportion of men and women in Bogotá who support maternal employment. While there are no statistically significant changes in these aggregate perceptions, the gender gap in perceived support remains evident, particularly with respect to beliefs about women.

Columns (4) and (5) focus on misperception indicators —binary variables equal to one if the respondent underestimates true support levels. Since treatment effects vary by gender, we interpret the coefficients using both the main treatment effect (β_1) — which captures the effect for men— and the interaction term (β_3), which reflects the difference in treatment response between women and men. The effect for women is given by the sum $\beta_1 + \beta_3$. For transparency, we also report regression estimates for the subsample of women, which yield nearly identical results.[1] In the female-only OLS estimates (Panel A, Sub-Sample Women), the treatment reduces misperceptions about other women's support by 7.2 percentage points (column 5), suggesting a reduction of approximately 13 percentage points relative to the baseline control mean (55.3%). This result is reinforced in the IPW estimates (Panel B), where the estimated average treatment effect on the treated (ATT) for women is a reduction of 7.3 percentage points. These consistent negative effects across specifications support the interpretation that the intervention helped women correct over-pessimistic beliefs about other women's attitudes.

Among men, the point estimate in column 5 (Panel A) is positive —suggesting a possible increase in misperceptions about women's beliefs— but this effect is small,

[1]Formally, the treatment effect for women is $\beta_1 + \beta_3$. The estimates from the female-only specification are numerically very similar.

imprecisely estimated, and not robust to the IPW specification.

Taken together, these results indicate that the intervention selectively improved belief accuracy —reducing women's misperceptions about peer norms, while having no impact on their own attitudes or beliefs about men's support. These asymmetries are consistent with domain- and gender-specific belief updating, and suggest that information is most likely to shift perceptions where initial misalignment is large and personally salient. These findings are consistent with an intensive-margin mechanism: rather than shifting average attitudes, the intervention corrected belief errors among a subset of respondents whose prior beliefs were most misaligned with reality.

Table 7. Beliefs and misperceptions about beliefs of other men and women in Bogota regarding working mothers (Sample 2 ∩ 3)

	First-Order Belief	Second-Order Beliefs		Misperceptions	
	(1)	(2) Men	(3) Women	(4) Men (D)	(5) Women (D)
Panel A: OLS estimates, Sample 2 ∩ 3 (WhatsApp and Endline)					
Treatment (β_1)	002	.237	–.846	–.033	.087**
{Std. Err}	(.045)	(2.790)	(2.833)	(.057)	(.066)
Fisher p-value	[0.916]	[0.881]	[0.601]	[0.264]	[0.028]
Female (β_2)	.039*	–3.962	–2.769***	.055	.113***
{Std. Err}	(.037)	(2.388)	(2.411)	(.045)	(.059)
Fisher p-value	[0.068]	[0.946]	[0.000]	[1.000]	[0.000]
Treatment × Female (β_3)	.005	1.488	2.929***	.001	–.160***
{Std. Err}	(.053)	(3.361)	(3.379)	(.065)	(.079)
Fisher p-value	[0.985]	[0.995]	[0.000]	[0.998]	[0.000]
Constant	.784***	48.958***	49.874***	.871***	1.140***
	(.109)	(7.947)	(8.299)	(.151)	(.196)
N	689	689	689	689	689
Mean Dep. Var	0.907	64.907	78.907	0.801	0.511
OLS estimates, Sample 2 ∩ 3: Subsample of Women					
Treatment	.009	1.727	2.043	–.035	–.072*
{Std. Err}	(.025)	(2.057)	(1.943)	(.037)	(.047)
Fisher p-value	[0.720]	[0.384]	[0.358]	[0.320]	[0.102]
Panel B: IPW estimates, Sample 2 ∩ 3 (WhatsApp and Endline)					
ATT Male	.011	.456	.076	–.026	.084
B.S. Std. Err	{0.057}	{3.286}	{3.281}	{.076 }	{0.079}
p-value	.847	.890	.982	.730	.288
Mean Dep. Var (Controls)	0.864	67.785	79.60	0.758	0.434
ATT Female	–.001	1.192	2.579	–.022	–.073*
Bs Std. Err	{0.028}	{2.253}	{2.311}	{0.040}	{0.041}
p-value	.962	.597	.264	.580	.075
Mean Dep. Var (Controls)	.922	63.632	76.774	.824	.553

Notes: Standard errors in parentheses ( ); Fisher exact p-values in brackets [ ]; Bootstrapped (Bs) standard errors in IPW results are in {}. Stars reflect Fisher p-values for regression outcomes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All OLS models include the following controls: socio-demographic covariates (age, income), baseline outcomes, and stratification variables (e.g., wife's employment status and baseline beliefs). Inverse Probability Weighting (IPW) estimates are clustered at the individual level, with gender-specific treatment effects (ATT) reported separately. Results in the subsample of women reflect regressions restricted to female respondents. Treatment effects for women in the full sample correspond to the sum $\beta_1 + \beta_3$; these are numerically similar to the female-only estimates.

Second-order beliefs about spouse, Y_s. Next, we examine whether the intervention affected participants' beliefs about their partner's views —testing for within-household

belief updating. The estimation sample remains Sample 2 ∩ 3, as in Table 7. However, this analysis captures a *reinforced treatment*: treated respondents were re-exposed to the information immediately before reporting their beliefs about their partner's views during the endline phone interview.

Table 8 presents results for two different statements: support for maternal employment (columns 1–3) and equal sharing of household tasks (columns 4–6). For each statement, we report treatment effects on three different outcomes: (i) respondents' own attitudes (first-order beliefs), (ii) perceived spousal support (second-order beliefs), and (iii) binary misperception indicators equal to one if the respondent underestimates their partner's support.

We find no effect on own attitudes (columns 1 and 4), which were already high at baseline (91% for maternal employment and 98.5% for task sharing). In contrast, the intervention significantly improved perceptions of the spouse's attitudes —particularly among men— and reduced misperceptions.

For *maternal employment*, the main treatment coefficient (β_1) indicates that treated *men* increased their perceived partner support by 7.3 percentage points (column 2), and were 15.9 percentage points less likely to misperceive it (column 3). The IPW-adjusted estimates confirm these effects: the average treatment effect on the treated (ATT) for men is 9.1 percentage points for perceived support and –18.1 percentage points for misperception, representing a 42% reduction from the baseline misperception rate (43%).

For *women*, the treatment effect on perceived spousal support is estimated directly from the subsample regression: an increase of 5.4 percentage points (p = 0.046), consistent with the implied sum $\beta_1 + \beta_3$.[2] However, there is no statistically significant effect on misperceptions. IPW estimates corroborate this pattern: ATT estimates for women

[2]The sum of coefficients $\beta_1 + \beta_3$ approximates the treatment effect for women and closely matches the estimate from the subsample regression.

show a 5.9–6.3 percentage point increase in perceived support across both domains, but no detectable change in misperception.

For *equal task sharing*, we again find no change in own attitudes (columns 4 and 5), which were near-universal. But the intervention significantly increased perceived partner support: by 6.6 percentage points among women (subsample estimate), and by 9.1 percentage points among men (main effect β_1). The reduction in misperceptions is large and statistically significant among men—20 percentage points (column 6), or a 45% decline relative to the control mean (44.9%). Among women, the effect is negative but small and not statistically significant (–3.9 percentage points).

In summary, the intervention —though targeted at societal beliefs— elicited notable improvements in intra-household belief accuracy, particularly among men and in domains where misperceptions were more prevalent. These effects occurred *on the intensive margin*, without altering respondents' own attitudes. The evidence is consistent with a *generalized introspection mechanism*, whereby exposure to corrected normative beliefs prompts reassessment of assumptions about close others.

Table 8. Beliefs and misperceptions about spouse's attitudes: Working mothers and equal task sharing (Sample 2 ∩ 3)

	Working Mothers			Equal Task Sharing		
	First-order (1)	Second-order (Spouse) (2)	Misperception (D) (3)	First-order (4)	Second-order (Spouse) (5)	Misperception (D) (6)
Panel A: OLS estimates, Sample 2 ∩ 3 (WhatsApp and Endline)						
Treatment (β_1)	.002	.073***	−.159***	.004	−.012	−.163***
(Std. Err.)	(.045)	(.033)	(.060)	(.022)	(.018)	(.060)
[Fisher p-value]	[.916]	[.000]	[.000]	[.653]	[.542]	[.000]
Female (β_2)	.039*	−.012	.067	.007***	−.100***	.068
(Std. Err.)	(.037)	(.034)	(.052)	(.017)	(.023)	(.053)
[Fisher p-value]	[.068]	[.999]	[.999]	[.003]	[.000]	[.999]
Treatment x Female (β_3)	.005	−.025	.128***	.019	.078***	.135***
(Std. Err.)	(.053)	(.042)	(.072)	(.024)	(.033)	(.072)
[Fisher p-value]	[.998]	[1.00]	[.000]	[1.000]	[0.00]	[0.00]
Constant	.784***	.836***	.686***	.988***	.577***	.746***
(Std. Err.)	(.109)	(.109)	(.186)	(.042)	(.113)	(.185)
N	689	689	689	689	689	689
Mean Dep. Var	.906	.922	.434	.985	.907	.434
OLS estimates, Sample 2 ∩ 3 (Subsample Women)						
Treatment	.009	.054**	−.029	.023	.066**	−.030**
(Std. Err.)	(.025)	(.027)	(.048)	(.010)	(.030)	(.048)
Fisher p-value	[.720]	[.046]	[.553]	[.012]	[.035]	[.529]
Panel B: IPW estimates, Sample 2 ∩ 3 (WhatsApp and Endline)						
ATT Male	.011	.091**	−.181**	.040	−.006	−.199***
{Bs Std. Err}	{.057}	{.043}	{.072}	{.041}	{.019}	{.075}
p-value	[.847]	[.035]	[.012]	[.321]	[.740]	[.008]
Mean Dep. Var (Controls)	.864	.882	.431	.933	.989	.449
ATT Female	−.001	.059*	−.037	.026*	.063**	−.039
{Bs Std. Err}	{.028}	{.034}	{.053}	{.015}	{.032}	{.053}
p-value	[.962]	[.083]	[.485]	[.092]	[.046]	[.461]
Mean Dep. Var (Controls)	.922	.870	.513	.974	.840	.515

Notes: Standard errors in parentheses ( ); Fisher exact *p*-values in brackets [ ]; Bootstrapped (Bs) standard errors in IPW results are in {}. Stars reflect Fisher *p*-values for regression outcomes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All OLS models include the following controls: socio-demographic covariates (age, income), baseline outcomes, and stratification variables (e.g., wife's employment status and baseline beliefs). Inverse Probability Weighting (IPW) estimates are clustered at the individual level, with gender-specific treatment effects (ATT) reported separately. Results in the subsample of women reflect regressions restricted to female respondents. Treatment effects for women in the full sample correspond to the sum $\beta_1 + \beta_3$; these are numerically similar to the female-only estimates.

Changes in Intra-Household Preferences for Course Enrollment

We next assess whether the intervention influenced participants' stated preferences related to women's labor force participation (LFP) within the household. Specifically, we examine whether treated individuals were more likely to (i) select the wife to enroll in a career-enhancing course, (ii) express personal interest in enrolling themselves, or (iii) believe that their partner is interested in participating. These outcomes —captured directly through the WhatsApp interface— are shown in columns (1) to (3) of Table 9. Unlike prior outcomes, the estimation sample here includes all individuals who engaged with the chatbot, i.e., sample (2), regardless of whether they later completed the endline phone survey. Therefore, we have a larger sample relative to the beliefs outcomes.

Panel A presents OLS estimates controlling for baseline and stratification variables. Panel B shows IPW-adjusted treatment effects separately by gender. As noted earlier, the main effect (β_1) corresponds to the treatment effect for male respondents, while $\beta_1 + \beta_3$ gives the effect for women. In column (1) of Table 9, however, the outcome differs by gender: for men, it captures willingness to prioritize their wife; for women, it captures willingness to prioritize themselves. The construction thus reflects prioritization of the wife within the couple.

Among men, the intervention significantly increased the likelihood of selecting their wife for the course: the estimated effect is 7.5 percentage points (p = 0.004), representing a 23% increase relative to the male control mean of 32.7 percent. The IPW-adjusted ATT for men is 8.8 percentage points (significant at 10 percent), consistent in magnitude. These findings suggest that the intervention successfully increased men's willingness to prioritize their wife for a career-building opportunity.

Among women, however, the decision is self-regarding. The control group baseline is already very high—around 75%—leaving limited scope for upward adjustment. Con-

sistent with this ceiling, we detect no treatment effect: the estimate for women is close to zero and statistically insignificant across specifications. Indeed, the sum of β_1 and β_3 remains near zero, indicating no additional effect on female participants, when we restrict the sample to women results remain, and the IPW are also not different from zero.

Columns (2) and (3) examine own interest and beliefs about partner interest. We find no statistically significant effects for either outcome, in either gender. Men's and women's willingness to attend the course themselves, and their beliefs about their partner's interest, remain unchanged. This reinforces the view that the intervention shifted behavior where initial priors were malleable and potentially misaligned —namely, men's views about whether their wife should be prioritized.

Overall, these results align with our broader interpretation: the intervention operates on the intensive margin, primarily correcting misperceptions rather than transforming already-formed attitudes. Where preferences were strong or well-informed —especially among women— additional information had no measurable impact.

Table 9. Changes in interest about wife's LFP, Sample 2

Dependent Variable	Wife Should Attend Course	Self: Interested	Partner: Believed Interested
	(1)	**(2)**	**(3)**
Panel A: OLS estimates, Sample 2 (WhatsApp Survey)			
Treatment β_1	.076***	−.030	.013
{Std. Err}	(.051)	(.047)	(.052)
Fisher p-value	[0.005]	[0.261]	[0.684]
Female β_2	.429***	.105	−.241
{Std. Err}	(.042)	(.038)	(.046)
Fisher p-value	[0.000]	[1.000]	[0.999]
Treatment $\times$ **Female** β_3	−.087***	.028***	−.018***
{Std. Err}	(.058)	(.054)	(.063)
Fisher p-value	[0.000]	[0.000]	[0.000]
Constant	.327**	.437**	.742***
	(.169)	(.171)	(.186)
N	1019	986	975
Mean Dep. Var	.706	.789	.461
OLS estimates, Sample 2 (Subsample Women)			
Treatment	−.008	−.001	−.000
{Std. Err}	(.029)	(.031)	(.039)
Fisher p-value	[0.783]	[0.980]	[0.989]
Panel B: IPW Estimates, Sample 2 (WhatsApp Survey)			
ATT Male	.088*	−.026	.017
{B.S. Std. Err}	{0.053}	{0.066}	{0.073}
p-value	[0.101]	[0.698]	[0.811]
Mean Dep. Var (Controls)	.466	.766	.643
ATT Female	−.014	.023	.025
{B.S. Std. Err}	{0.039}	{0.039}	{0.052}
p-value	[0.711]	[0.553]	[0.622]
Mean Dep. Var (Controls)	.844	.837	.386

Notes: Standard errors in parentheses ( ); Fisher exact p-values in brackets [ ]; Bootstrapped (Bs) standard errors in IPW results are in {}. Stars reflect Fisher p-values for regression outcomes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All OLS models include the following controls: socio-demographic covariates (age, income), baseline outcomes, and stratification variables (e.g., wife's employment status and baseline beliefs). Inverse Probability Weighting (IPW) estimates are clustered at the individual level, with gender-specific treatment effects (ATT) reported separately. Results in the subsample of women reflect regressions restricted to female respondents. Treatment effects for women in the full sample correspond to the sum $\beta_1 + \beta_3$; these are numerically similar to the female-only estimates.

Labor Market Actions and Attitudes

We conclude the empirical analysis by testing whether the intervention affected labor market behaviors or stated preferences related to work. Table 10 presents treatment effects on three outcomes: (i) whether the respondent changed jobs or started a business in the past month, (ii) whether they express a desire to improve their labor market situation, and (iii) their stated preference for balancing work and family life. All outcomes were collected at endline.

Importantly, treatment exposure varied: some respondents were treated during the midline WhatsApp survey, others only at endline, and some were indirectly exposed through a treated partner. Because change in actual labor participation (outcome (i)) can plausibly be influenced only by information received prior to the relevant window of behavior, we use multiple samples to distinguish between individuals who were treated a few months before and those for whom the exposure to the intervention was too recent.

Columns (1), (4), and (6) restrict the sample to individuals who received the WhatsApp message and were later re-surveyed. These individuals were exposed twice: first via WhatsApp, and again through a scripted reminder administered immediately prior to endline questions. Columns (2), (5), and (7) use broader samples. Column (2) includes also individuals who did not receive the WhatsApp message themselves but lived with a treated partner and later completed the endline survey —thus identifying potential spillovers from within-household exposure. Columns (5) and (7) comprise all directly treated individuals, regardless of whether the information was received only at endline or also at midline via WhatsApp, and capture the effect of receiving at least one dose of treatment. Finally, column (3) isolates respondents who were treated only at endline, and thus the information nudge could not have affected labor behavior, then it provides a placebo for interpreting the job-change results (column (3)).

The clearest effects emerge in column (1), where we find that the intervention increased the likelihood that treated women changed jobs or started a business, by 8.5 percentage points, relative to a control mean of 75.2%. The corresponding IPW-adjusted ATT for women is of similar magnitude (8.0pp), and this pattern holds in the indirect exposure sample (column 2), where women not directly treated but living with a treated partner also exhibit higher job mobility (ATT = 10.8pp). No effects are observed among men or in the placebo sample (column 3), consistent with the interpretation that only initial WhatsApp exposure—delivered prior to the job-change window—could plausibly drive behavioral adjustments.

For the outcome on the aspiration to improve labor market outcomes (columns 4 and 5), estimated effects are small and imprecise across all subgroups, though point estimates among treated women are consistently positive. For the two attitudinal outcomes, the timing of measurement allows for possible reinforcement effects through repeated exposure. Among men, we find robust and statistically significant increases in the preference for balancing work and family life. In the double-exposure sample (column 6), the treatment effect is 8.4 percentage points; in the full endline sample (column 7), the effect remains positive and significant at 5.8 percentage points. IPW estimates confirm these findings, suggesting that normative updating through information prompts introspection around household role division. In contrast, we find no such effects among women, whose baseline support for equal task sharing was already high.

These findings underscore the value of the couple-based design, which allows us to separately identify effects by gender and explore intra-household spillovers. The clearest behavioral changes appear among women directly exposed to the intervention: in the sample that received both WhatsApp and endline reminders (column 1), treated women were 8.5 percentage points more likely to have changed jobs or initiated a business, a significant shift relative to an already high control mean. This effect is robust to inverse

probability weighting and replicates among women who were not directly treated but lived with a treated partner, indicating meaningful within-household spillovers.

More broadly, the couple-based design allows us to uncover asymmetries in how information affects beliefs and aspirations across genders—insights that would be masked in individual-level interventions. While we observe behavioral responses among women, a key result lies in the attitudinal shifts among men: repeated exposure to the intervention significantly increases their preference for balancing work and family life, a change that suggests greater openness to more equal sharing of domestic responsibilities. This evolution is far from trivial. As emphasized by Goldin (2021), progress toward gender equality in labor markets hinges not only on women's behavior, but also on men's willingness to shift traditional roles at home. Yet such changes in norms are often slow and difficult to achieve. That we observe a significant movement in men's stated aspirations is encouraging—and highlights the potential of light-touch, couple-targeted interventions to spark reflection. Future research should explore whether these attitudinal shifts endure over time and whether they eventually translate into more equal divisions of labor within the household.

Table 10. Employment Outcomes

	Started to change job			**Aspire better LM situation**		**Wants work and family balance**	
	Sample 2 ∩ 3	Ind. or Dir exp. (Sample 3)	Placebo	Sample 2 ∩ 3	Sample 3	Sample 2 ∩ 3	Sample 3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS estimates with baseline covariates							
Treatment β_1	-0.027	0.010	-0.066	-0.008	-0.011	0.084*	0.058**
{Std. Err}	(0.059)	(0.042)	(0.064)	(0.067)	(0.041)	(0.064)	(0.039)
Fisher p-value	[0.442]	[0.756]	[0.121]	[0.828]	[0.659]	[0.017]	[0.026]
Female (=1)	-0.017***	0.133***	0.050	-0.003	0.014	0.070***	0.043***
{Std. Err}	(0.047)	(0.042)	(0.055)	(0.059)	(0.039)	(0.056)	(0.037)
Fisher p-value	[0.000]	[0.000]	[1.000]	[0.061]	[1.000]	[0.000]	[0.000]
Treatment × Female	0.110***	0.082	0.016	0.057***	0.030***	-0.069***	-0.044***
{Std. Err}	(0.068)	(0.057)	(0.080)	(0.083)	(0.054)	(0.080)	(0.052)
Fisher p-value	[0.000]	[1.000]	[1.000]	[0.002]	[0.000]	[0.000]	[0.000]
Constant	.565***	.467***	.751***	.323*	.363***	.323*	.385***
	(.167)	(.142)	(.207)	(.194)	(.133)	(.194)	(.133)
N	689	1138	475	689	1388	689	1388
Mean Dep. Var	.752	.582	.714	.527	.500	.354	.365
OLS estimates with baseline covariates (Sub-Sample Women)							
Treatment	0.085**	0.098***	-0.044	0.052	0.024	0.018	0.012
{Std. Err}	0.038	0.039	0.051	0.047	0.036	0.046	0.035
Fisher p-value	.019	.009	.393	.282	.464	.673	.712
Panel B: Inverse Probability Weighting (IPW) estimations							
ATT Male	-0.034	0.016	-0.083	-0.016	-0.029	0.062	0.063*
{B.S. Std. Err}	.070	.043	.069	.078	.044	.075	.041
Mean Dep. Var (Controls)	.730	.473	.738	.516	.506	.322	.320
ATT Female	.080*	.108**	–.064	.053	.022	.014	.012
{B.S. Std. Err}	.042	.043	.053	.051	.039	.049	.037
Mean Dep. Var (Controls)	.735	.621	.798	.506	.499	.356	.368

Notes: Standard errors in parentheses ( ); Fisher exact p-values in brackets [ ]; Bootstrapped (Bs) standard errors in IPW results are in {}. Stars reflect Fisher p-values for regression outcomes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All OLS models include the following controls: socio-demographic covariates (age, income), baseline outcomes, and stratification variables (e.g., wife's employment status and baseline beliefs). Inverse Probability Weighting (IPW) estimates are clustered at the individual level, with gender-specific treatment effects (ATT) reported separately. Results in the subsample of women reflect regressions restricted to female respondents. Treatment effects for women in the full sample correspond to the sum $\beta_1 + \beta_3$; these are numerically similar to the female-only estimates. *Sample definitions by column:* Column (1): Sample 2 ∩ 3 — directly treated at midline and surveyed at endline. Column (2): Individuals in Sample 3 who either were directly treated or lived with a treated partner (i.e., indirect or direct exposure). Column (3): Placebo — surveyed at endline but not treated at midline; treated only at endline. Columns (4) and (6): Sample 2 ∩ 3 — directly treated at midline and surveyed at endline. Columns (5) and (7): Full Sample 3 — all individuals surveyed at endline, regardless of treatment timing.

6 Conclusion

This paper investigates whether pluralistic ignorance—systematic misperceptions of others' attitudes toward gender roles—contributes to persistent gender gaps in labor force participation. Using experimental data from cohabiting couples with young children in Bogotá, we document substantial second-order belief gaps. While nearly all individuals express support for maternal employment, both men and women significantly underestimate societal approval, particularly among men. These misperceptions extend into the household: women underestimate their partners' support for maternal employment, while men overestimate their wives' reluctance to share domestic responsibilities.

To test whether belief correction influences behavior, we implemented a randomized information intervention via WhatsApp. Treated participants received individualized feedback on prevailing local attitudes; controls received unrelated labor market content. Immediately afterward, all participants were asked to allocate a career-enhancing course within the household. This constrained decision allows us to test whether information reshapes intra-household prioritization under normative uncertainty.

The intervention improved belief accuracy, particularly within couples. While own attitudes and beliefs about broader societal norms remained unchanged, treated men significantly revised upward their perception of their partners' support, and treated women adjusted upward their beliefs about their partners' willingness to share caregiving. These intra-household revisions translated into meaningful behavioral responses: treated men were 7–8 percentage points (16%) more likely to prioritize their wives for the course, while women—whose baseline prioritization was already high—did not respond. Treated women also reported increased job search and aspirations to improve employment, and treated men expressed a stronger preference for work–family balance.

We estimate treatment effects using OLS with stratification fixed effects, complemented by exact Fisher tests and inverse probability weighting (IPW) to adjust for non-random exposure and differential attrition. Results are robust across specifications and subgroups. All outcome-specific analyses are based on clearly defined samples—distinguishing between respondents exposed at midline via WhatsApp, at endline via phone, or indirectly through a partner. The short follow-up period (1–3 months) implies that findings reflect short-run adjustments rather than persistent change.

Several limitations deserve emphasis. First, participant engagement with the digital intervention was partial: not all assigned participants viewed the information, and intensity of exposure likely varied. Second, follow-up data were collected shortly after the intervention, precluding medium- or long-term inference. Third, all outcomes are self-reported and include stated intentions as well as behaviors. Fourth, attrition between baseline and follow-up was non-trivial and systematically related to baseline characteristics—particularly gender, education, household structure, and prior beliefs. We incorporate a detailed attrition analysis in the main text and implement IPW to mitigate selection bias.

Finally, the intervention's scalability should be interpreted with caution. While WhatsApp-based delivery is inexpensive, constructing accurate normative feedback requires population-level diagnostics. That is, the intervention's low marginal cost depends on prior investments in representative survey data.

Despite these limitations, our results highlight how misperceived norms can constrain decision-making even in settings with broadly egalitarian preferences. Correcting belief distortions—particularly about close others—shifted within-household choices along relevant margins. These findings suggest that norm-based informational interventions, when paired with credible diagnostics, can complement structural labor market policies by addressing the social beliefs that shape household behavior.

References

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer-Moore, "Male Social Status and Women's Work," *AEA Papers and Proceedings*, May 2018, *108*, 363–67.

Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate statistical discrimination: An identification problem," *Review of Economics and Statistics*, 2023, pp. 1–45.

Boltz, Marie, Ana Maria Diaz, and Luz Magdalena Salas, "Hidden Biases: Pluralistic Ignorance in Perceptions of Women's Work Among Couples in Colombia," *Revue d'Economie du Développement*, 2025.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott, "Misperceived social norms: Women working outside the home in Saudi Arabia," *American economic review*, 2020, *110* (10), 2997–3029.

_ , Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott, "How are gender norms perceived?," Technical Report, National Bureau of Economic Research 2023.

Cameron, Lisa, Diana Stella Contreras Suarez, and Diahhadi Setyonaluri, "Leveraging Women's Views to Influence Gender Norms around Women Working: Evidence from an Online Intervention in Indonesia," Technical Report, The World Bank 2024.

Frisancho, Verónica, Monserrat Bustelo, and Mariana Viollaz, "What Is the Labor Market Like for Women in Latin America and the Caribbean?," 2023.

Goldin, Claudia, *Career and family: Women's century-long journey toward equity*, Princeton University Press, 2021.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 181–209.

Laszlo, Sonia, Carolina de Miranda, Franque Grimard, Tania Insfran, Thomas Kokossou, Laura Morinigo, and Jim Warnick, "Norm shifting interventions and beliefs about gender norms in Paraguay," 2023.

Appendix

Appendix A: Chatbot Flow for Treated Participants

Message 1: Introduction

Hi, [NAME]!

Thank you for participating in our survey with Universidad Javeriana and IPSOS about household characteristics in Bogotá.

Now, we'd like to invite you to discover the results and reflect on your answers. Your opinion continues to make a difference!

You may also receive a special resource.

Are you in? It will only take 5 minutes

Options: **Yes [Continue to Message 2] / No [Continue to Reminder]**

Random Block 1: Perceptions about Women

Message 2

Question: Out of 100 women with children under 6 in Bogotá, how many do you think believe mothers should have the freedom to work outside the home?

Message 3

Your answer was: [N] out of 100

Message 4

Do you think you were correct?

1. Yes

2. No

3. Not sure

Message 5

The reality was: [X] out of 100 women really support this idea

Message 6

What do you think about this difference?

1. Interesting

2. Irrelevant

3. Disappointing

Random Block 2: Perceptions about Men

Message 7

Question: Out of 100 men with children under 6 in Bogotá, how many do you think believe mothers should have the freedom to work outside the home?

Message 8

Your answer was: [N] out of 100

Message 9

Do you think you were correct?

1. Yes

2. No

3. Not sure

Message 10

Find out what Bogotá's fathers really think
X
out of 100 actually support this idea

Message 11

What do you think about this difference?

1. Interesting

2. Irrelevant

3. Disappointing

Final Block: Course and Household Decisions

Message 12

As a thank-you for participating, we offer you an online course to develop tools for accessing new job opportunities and improving your job search

Message 13

If there are more interested people than spots, we may assign one per household.

Who do you think should participate?

1. Me

2. My partner

Message 14

Would you be interested in taking the course?

1. Yes

2. No

Message 15

Do you think your partner would be interested in the course?

1. Yes

2. No

3. I don't know

Message 16

We value your opinion! We'd like to ask two more questions based on hypothetical scenarios.

Message 17

A woman has the chance to accept a well-paid job . To do so, she must pay for childcare, which would leave the family income unchanged

What should she do?

Accept the job

Decline the job

Message 18

Alternatively, instead of paying for childcare, her husband could work more from home to care for the child, but he would earn less, again leaving household income unchanged

What should the woman do?

Accept the job

Decline the job

Closing Message

Thank you for participating! We'll call you soon with a few follow-up questions and more information about the course for you or your partner. Stay tuned!

Appendix B: Chatbot Flow for Control Group Participants

Message 1: Introduction

Thank you for participating in our survey with Universidad Javeriana and IPSOS about household characteristics in Bogota.

Now, we would like to invite you to discover the results and reflect on your answers. Your opinion continues to make a difference.

You may also receive a special resource.

Are you in? It will only take 5 minutes.

Options: Yes [Continue to Message 2] / No [Go to Reminder]

Reminder

We understand you might be short on time, but your participation is very valuable to us.

We would love for you to see the survey results and what they reveal about households in Bogota. It is a chance to compare your responses with others and learn more.

Options: Yes, I will participate [Continue to Message 2] / No, I will not continue [Go to Exit Question]

Block 1: Social Perception on Environmental Policy

Message 2

Do you think companies should subsidize employees who use public transport or car-pool to reduce commute-related emissions?

Message 3

Your answer was: [Agree / Disagree]

Participants were then randomly assigned to receive perceptions about women or men.

Message 4

Do you think women in Bogota share your opinion?

1. Yes

2. No

3. I do not know

Message 5

X

out of 100 women with children under 6 in Bogota support this idea.

Message 6

What do you think of this information?

1. Interesting

2. Irrelevant

3. Disappointing

Message 7

Do you think men in Bogota share your opinion?

1. Yes

2. No

3. I do not know

Message 8

X

out of 100 men with children under 6 in Bogota support this idea.

Message 9

What do you think of this information?

1. Interesting

2. Irrelevant

3. Disappointing

Final Block: Course and Household Decisions

Message 10

As a thank-you for participating, we offer you an online course to develop tools for accessing new job opportunities and improving your job search.

Message 11

If there are more interested people than spots, we may assign one per household.

Who do you think should participate?

1. Me

2. My partner

Message 12

Would you be interested in taking the course?

1. Yes

2. No

Message 13

Do you think your partner would be interested in the course?

1. Yes

2. No

3. I do not know

Message 14

We value your opinion. We would like to ask two more questions based on hypothetical situations.

Message 15

A woman has the opportunity to accept a well-paid job. To do so, she would need to pay for childcare, keeping household income unchanged.

What should she do?

1. Accept the job

2. Decline the job

Message 16

Alternatively, her husband could work more from home to care for the child, earning less, and keeping household income unchanged.

What should the woman do?

1. Accept the job

2. Decline the job

Closing Message

Thank you for participating. In the coming days, we will contact you with some follow-up questions and more information about the course for you or your partner.

Table 11. Balance Test for Different Samples (Dependent Variable: Treated = 1)

Variable	Midline Survey *Sample (2)*	Endline Survey *Sample (3)*	Both Surveys *Sample (2)∩(3)*	Endline Only (Excl. Mid) *Sample (3)\(2)*
Female (=1)	0.011	0.146	0.076	0.241
	(0.137)	(0.113)	(0.169)	(0.162)
Household members	0.028	0.126*	0.109	0.183*
	(0.081)	(0.071)	(0.099)	(0.105)
Adults 18–60 years old	−0.175	−0.104	−0.182	−0.102
	(0.139)	(0.117)	(0.174)	(0.165)
Adults older than 60 years	0.196	0.211	0.231	0.096
	(0.242)	(0.211)	(0.313)	(0.295)
Permanent care (1 = Yes)	0.133	0.024	0.169	−0.119
	(0.135)	(0.117)	(0.167)	(0.168)
Age	0.011	0.012*	0.010	0.015
	(0.009)	(0.007)	(0.011)	(0.010)
Low education	−0.233	0.020	−0.256	0.219
	(0.277)	(0.223)	(0.349)	(0.302)
Medium education	−0.054	−0.068	−0.184	0.037
	(0.184)	(0.153)	(0.222)	(0.217)
Unemployed	−0.340	−0.141	−0.311	0.044
	(0.269)	(0.233)	(0.336)	(0.333)
Stratum 1	−0.176	−0.030	−0.085	−0.016
	(0.341)	(0.301)	(0.446)	(0.414)
Stratum 2	0.168	−0.005	0.200	−0.206
	(0.289)	(0.252)	(0.370)	(0.352)
Stratum 3	−0.002	0.130	0.122	0.192
	(0.289)	(0.253)	(0.365)	(0.360)
Stratum 4	0.528	0.405	0.678	0.207
	(0.354)	(0.294)	(0.450)	(0.398)
First-order belief: Women work	−0.166	0.099	−0.498	0.608
	(0.505)	(0.382)	(0.669)	(0.504)
SOB (men): Women work	0.003	0.005	0.005	0.003
	(0.004)	(0.003)	(0.005)	(0.005)
SOB (men, SDB): Women work	−0.004	−0.005	−0.006	−0.003
	(0.004)	(0.003)	(0.004)	(0.004)
SOB (women): Women work	0.006	0.003	0.008*	−0.002
	(0.004)	(0.003)	(0.005)	(0.005)
FOB: Moms under 6 work	−0.145	−0.386*	−0.333	−0.458
	(0.220)	(0.199)	(0.279)	(0.298)
SOB (men): Moms under 6 work	0.003	0.000	0.002	−0.001
	(0.004)	(0.003)	(0.005)	(0.005)
SOB (women): Moms under 6 work	−0.000	−0.005	−0.003	−0.009*
	(0.004)	(0.003)	(0.005)	(0.005)
SOB (partner): Moms under 6 work	−0.432*	−0.391*	−0.587**	−0.203
	(0.232)	(0.216)	(0.298)	(0.326)
Constant	−0.043	−0.037	0.314	−0.209
	(0.780)	(0.622)	(1.005)	(0.847)
Observations	1,019	1,388	689	699

Notes: Each column reports coefficients from a separate OLS regression in which the dependent variable equals 1 if the individual was treated. All regressions include stratification controls. Standard errors, clustered at the household level, appear in parentheses. FOB = First-order belief; SOB = Second-order belief; SDB = Social-desirability bias. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Selective Attrition

To examine whether attrition was systematically related to baseline characteristics, we estimated separate logit models for each final sample. In every specification the dependent variable equals one if the respondent appears in the relevant sample and zero otherwise, while the covariates comprise the full set of pre-treatment individual and household variables. Table 12 reports the coefficients that are statistically different from zero at the 10% level in at least one model, thereby highlighting the principal correlates of study retention.

The estimates reveal clear evidence of non-random attrition. First, demographic variables matter: women were substantially more likely to be retained (e.g. a coefficient of 0.826 in the midline survey and 0.914 in the balanced panel), as were respondents living in larger households (coefficients between 0.212 and 0.233) and those reporting a household member in permanent care (coefficients between 0.210 and 0.282). Age is also positively associated with inclusion; the log-odds of remaining in the study rise by roughly 0.02 for each additional year. Conversely, households with fewer working-age adults or with elderly members were somewhat less likely to stay, as indicated by negative coefficients on the number of adults aged 18–60 and on adults older than 60 in the balanced sample.

Second, attrition skews the final samples toward higher socioeconomic status. Respondents from lower urban strata were systematically less likely to be retained: the log-odds of inclusion for Stratum 1 fall by as much as 0.827 in the endline model, with similarly large negative estimates for Strata 2–4. Educational differences are smaller, although individuals with medium education show a weakly positive association with completion in the midline survey, while low education is never significantly related to retention once other factors are controlled.

Third, we detect meaningful attitudinal selection. Respondents who endorsed the statement that "women work outside the home" (FOB1) were more likely to remain in the study, whereas stronger social-norm pressure against mothers of young children working (SOB2M) reduced the probability of appearing in the balanced panel. These patterns imply that attrition shifted the sample toward individuals with more progressive views on female labour supply.

Not all baseline covariates display systematic effects: the number of children, unemployment status, several second-order belief measures, and the high-education indicator are generally unrelated to study retention. Nevertheless, the significant coefficients documented above confirm that attrition was selective. The respondents who completed the surveys are disproportionately female, older, drawn from larger and better-off households, and hold more egalitarian gender attitudes. To mitigate the potential bias introduced by this non-random attrition, we use inverse probability weights derived from the logit models and includes the full set of baseline covariates as controls.

Table 12. Correlates of Attrition Across Sample Definitions (Logit Coefficients)

	Midline Survey *Sample (2)*	**Endline Survey** *Sample (3)*	**Midline & Endline** *Sample (2)∩(3)*	**Endline Only (No Mid)** *Sample (3)\(2)*
Female (=1)	0.826***	0.491***	0.914***	–0.143
	(0.081)	(0.073)	(0.094)	(0.088)
Household members	0.212***	0.233***	0.232***	0.109**
	(0.050)	(0.047)	(0.056)	(0.055)
Adults 18–60 years	–0.108	–0.109	–0.179*	0.009
	(0.083)	(0.078)	(0.096)	(0.091)
Adults older than 60 years	–0.176	–0.215	–0.334*	0.007
	(0.148)	(0.138)	(0.171)	(0.161)
Permanent care (1 = Yes)	0.262***	0.210***	0.282***	0.034
	(0.082)	(0.076)	(0.092)	(0.091)
Age	0.018***	0.017***	0.024***	0.003
	(0.005)	(0.005)	(0.006)	(0.005)
Low education	–0.190	–0.072	–0.286	0.120
	(0.162)	(0.143)	(0.186)	(0.168)
Medium education	0.198*	0.003	0.064	–0.059
	(0.111)	(0.101)	(0.124)	(0.121)
Unemployed	0.096	0.075	0.052	0.060
	(0.163)	(0.153)	(0.186)	(0.183)
Stratum 1	–0.655***	–0.827***	–0.584**	–0.609***
	(0.217)	(0.206)	(0.250)	(0.235)
Stratum 2	–0.627***	–0.628***	–0.382*	–0.494**
	(0.188)	(0.180)	(0.213)	(0.202)
Stratum 3	–0.426**	–0.454**	–0.038	–0.585***
	(0.189)	(0.181)	(0.212)	(0.204)
Stratum 4	–0.734***	–0.545***	–0.597**	–0.209
	(0.219)	(0.204)	(0.251)	(0.227)
FOB1: Women work outside the home	0.504*	0.201	0.483	–0.124
	(0.276)	(0.235)	(0.330)	(0.272)
SOB1H: Women work outside the home	–0.000	–0.002	–0.002	–0.001
	(0.002)	(0.002)	(0.003)	(0.003)
SOB1HSDB: Women work outside the home	–0.001	–0.000	–0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
SOB1M: Women work outside the home	0.003	0.003	0.003	0.001
	(0.002)	(0.002)	(0.003)	(0.003)
FOB2: Moms <6 work outside home	–0.235*	–0.033	–0.244	0.190
	(0.133)	(0.126)	(0.151)	(0.158)
SOB2H: Moms <6 work (men)	0.002	0.002	0.005	–0.001
	(0.002)	(0.002)	(0.003)	(0.003)
SOB2M: Moms <6 work (women)	–0.005*	–0.003	–0.005*	–0.000
	(0.002)	(0.002)	(0.003)	(0.003)
SOB2P: Moms <6 work (partner)	–0.039	0.128	0.041	0.152
	(0.141)	(0.135)	(0.161)	(0.169)
Constant	–2.313***	–1.620***	–3.194***	–1.608***
	(0.444)	(0.402)	(0.517)	(0.471)
N	3 464	3 464	3 464	3 464

Notes: Logit coefficients (standard errors clustered at the household level in parentheses). Each regression includes stratification variables. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.