How Social Norms Shape Behavior Change: An Experimental Study of Female Smartphone Adoption in Rural India

Giorgia Barboni, Erica Field, Charity Moore,
Rohini Pande, Natalia Rigol, Simone Schaner*

December 10, 2024

Abstract

This study investigates the impact of a large-scale intervention aimed at increasing female smartphone adoption and engagement in rural India, focusing on the interplay of digital skills training and community norms. We evaluate a randomized controlled trial of an intervention delivered in 212 villages that provided training to improve digital literacy and incentives to encourage female phone adoption. Results indicate that encouraging phone use results in significant improvements in women's smartphone ownership, usage, social interactions, and mental health, and a liberalization of beliefs about gender-appropriate phone use. However, paradoxically, all of these gains are significantly *lower* in villages randomly assigned to target a majority rather than a minority of women. In high saturation villages, not only is female phone use lower, but we observe direct evidence of male backlash, suggesting resistance to shifting gender norms. The study underscores the importance of considering community dynamics when designing female empowerment initiatives and contributes to the broader understanding of social norm change and its potential unintended consequences.

*Barboni: University of Warwick (email: giorgia.barboni@wbs.ac.uk) ; Field: Duke University (email: emf23@duke.edu); Pande: Yale University (rohini.pande@yale.edu); Rigol: Harvard Business School (email: nrigol@hbs.edu); Schaner: University of Southern schaner@usc.edu); Troyer Moore: The Ohio State University (email: moore.1736@osu.edu). We thank Hema Balarama, Anwesha Bhattacharya, Harsh Jaiswal, Harnoor Kaur, Fatema Patel, Tejita Tiwari, Tanya Vaidya, and Balasai Vanukuri for excellent data collection, processing and analysis support. This project was supported with generous support from the Bill and Melinda Gates Foundation, USAID DIV, The Cash Transfers for Child Health (CaTCH), the National Science Foundation, IWWAGE, JPAL Innovation in Government Initiative, the Weiss Family Fund, WEE-DiFine, and GLM.

1 Introduction

The majority of the world's population now owns a smartphone, a transition that occurred faster than predicted in part due to the Covid-19 pandemic (GSMA 2023). This expansion was mostly driven by low- and middle-income countries (LMICs); in India, for example, smartphone adoption grew from 21% in 2015 to 77% in 2022, and 99% of adults in India now live within the footprint of mobile broadband network (GSMA India Report, 2022). Yet these trends mask deep gender disparities in smartphone access and use. Despite early evidence that the gender gap, or the difference in smartphone access and use between men and women, was shrinking in India during the first year of the pandemic, it has since remained steady around 40% in both 2021 and 2022 (GSMA Mobile Gender Gap Report 2023). Existing research suggests closing gender gaps in access to information technologies, such as smartphones, will benefit women. Research shows that even basic 2G mobile phones improve market functioning and help people learn about job opportunities. Services delivered through phones can improve behavior across a range of domains including health, education, and finance (see Barboni et al. (2018) for a review).

An important question is whether gender disparities in technology adoption in this context are driven primarily by gender differences in skill, gender differences in returns to phone use, or by social norms that differentially constrain female access to communications technologies. To shed light on this question, this paper explores whether providing women with phone training and norms-compatible uses encourages female smartphone adoption, and examines whether adoption is sensitive to the number of women in the community that are encouraged to adopt. The experiment leverages a large phone distribution program implemented by the state of Chhattisgarh in 2018. The program, called Sanchaar Kranti Yojana, or SKY, distributed over 2 million free smartphones to women, along with free mobile data. To ensure phone recipients could use mobile internet, the government also constructed LTE (4G) towers to cover eligible communities that lacked high speed data coverage.

To understand the influence of social norms on smartphone adoption in this setting we implemented a randomized control trial in 212 program villages that provided basic training to SKY-eligible women designed to help them build the skills needed to engage with smartphones. At the individual-level, we vary whether women receive only the training, the training plus a phone service called Mor Awaaz, or the training, the service and an incentive for engaging with the service. The Mor Awaaz phone service involves weekly push calls with pre-recorded information on good health practices, government health services, and updates on India's workfare program, and is intended to build a use case that is in line with the prevailing norm that women should care for the household. The service also includes

monthly "pull calls", where trained enumerators call women to conduct a short survey to measure women's awareness of, and access to, public health and nutrition-linked services. The incentives intervention offers women an in-kind or cash payment to engage with pull calls.

To evaluate the influence of social norms on smartphone adoption, we also randomly vary at the village level the proportion of women in the village that are offered training - none (32 villages), a low level of saturation in which a minority of women are invited ti participate (90 villages), or a high level of saturation in which the majority of women are invited to participate (90 villages).

First, we find that the treatment lead to important improvements in women phone interactions as well as downstream impacts. Relative to the control group, the pooled treatment significantly improves female phone ownership and usage. Women in the treatment group are more likely to own their own smartphones and they score 0.15 standard deviation units higher on indexes of basic and advanced phone usage. We document the number of interactions that women have with different categories of social groups - relatives outside the household, friends, and Self-Help Group (SHG) members in a typical month and how many of those interactions happen over the phone. We find that treatment women score 0.11 standard deviation units higher on an index of phone interactions. Treated women talk with an extra 0.9 people per month over the phone – primarily driven by interactions with relatives outside the home – a 16% increase when compared to the control group.

Women assigned to training also experience important reductions in the incidence of depression. We collect PHQ-4 mental health questions and construct an index of the incidence of anxiety and depression. Treated women see a 0.346 point, or 5%, improvement in the incidence of anxiety and depression. We also measure self-esteem using the Rosenberg scale and construct a spousal emotional control index, but observe no significant differences on these dimensions.

Alongside changing actual phone engagement, treatment leads to a liberalization of women's beliefs around the appropriateness of women's phone use. Treated women score 0.16 standard deviation units higher on an index of beliefs, increasing the proportion of women who believe that it is appropriate for unmarried and married women to use phones with supervision. This liberalization in beliefs is not present among husband, who are already less liberal than their spouses in the control group. So the intervention leads to a widening in the gender gap in views about the appropriateness of female phone use.

Second, we find that while the Mor Awaaz service is popular with respondents, neither the service nor the incentives to participate amplify treatment effects over and above the training. Over the course of the intervention, we generated 10 pull calls and over 250 push

calls. Respondents complete 74% of pull call surveys and pick up 65% of push calls and these rates remain basically constant throughout the 3 years of intervention. In our endline survey, women assigned to Mor Awaaz or to incentives are 20% more likely to own a smart phone than women in the control group. But neither the service alone nor the added incentives lead to important gains in outcomes over and above training alone for other outcomes.

Third, we show that treatment effects tend to be larger under the low saturation treatment. While the high saturation treatment has economically important effects on phone ownership and usage, women in the low saturation treatment enjoy impacts that are between 63% and 150% larger than those in high saturation. We also observe differences in treatment effects between low and high saturation villages on men's norms (this is super weak) and women's social connections.

This pattern is difficult to reconcile with a simple story of norms change, in which we would anticipate that treating a higher fraction of women in the village encourages female adoption. For instance, women might update their expectations about the acceptability of phone use knowing that many other women around them were also encouraged to use phones. It is also inconsistent with substantial network effects, through which the value of phone use rises with the number of social connections that adopt. Instead, the pattern suggests a more complex model of norms influence on behavior, such as a tipping point model, in which a sufficiently high rate of adoption triggers a behavioral response within the community that acts to discourage phone use.

One likely explanation for such a response is backlash arising from untargeted members of the community, the vast majority of which are men, who are more likely to become aware of or feel threatened by married women's phone use as more in their community are encouraged to adopt.[1] Male backlash in this context is reasonable to anticipate given evidence from other settings of negative responses from men to interventions or policies that shift women's autonomy or earnings power. Broadly speaking, backlash can be modeled as an attempt to stop or slow the pace of social change that goes against one's ideological position, as in Wheaton (2022). If backlash is costly to enact, or if it responds more when the pace of change is faster, we may observe backlash kicking in only when behavior change is widespread enough - or, in the context of our intervention, at a high enough level of saturation.[2]

[1] Of course, it is also possible that backlash originates among women in nontargeted demographic groups who fear being left behind. For instance, untargeted women may struggle more to adopt smartphones so feel threatened by a new norm embracing this form of communication. Likewise, backlash could originate from specific strong-minded women in the target group who are more likely to be treated themselves and therefore become aware of the push in high treatment density villages.

[2] Other models of backlash would also explain the pattern. For instance, backlash could arise in response to a low probability event such as a singular case of severe harassment that occurs with higher probability

Leveraging random variation in the proportion of a village that is treated, we find that women's phone engagement follows a parabolic pattern: phone use increases sharply from the control group level of 0 up to an apparent tipping point of around 40% of village treated, and then declines just as steadily back to zero by the time saturation hits 80%. Consistent with a theory of male backlash, while female norms become more liberal as a larger proportion of a village is treated, men's views about the appropriateness of women's phone use becomes more conservative through the saturation distribution.

As an alternative piece of evidence on backlash, we also exploit variation in the degree of baseline resistance to female phone use, which varies across households in our sample. We estimate the relationship between husband's baseline norms and their wives endline phone use separately for low saturation, high saturation, and control households. The estimates indicate a stronger relationship between phone use and norms in high saturation relative to low saturation areas, consistent with backlash. In other words, high saturation appears to make conservative men more conservative and progressive men more progressive with respect to their wives' phone use. Female phone use is lower in high saturation areas relative to low saturation areas only among men with below median baseline levels of acceptance. Among relatively more progressive men, levels of phone use in high saturation areas exceed levels of phone use in low saturation areas, suggesting that saturation magnifies norms of acceptability.

Our work speaks to three gaps in the literature. First, while extensive work has focused on the social and economic impacts of mobile phones, including basic and feature phones, little exists on smartphones. Basic mobile phones have been shown to improve individual well-being and promote productivity by enhancing market functioning (Jensen, 2007; Aker, 2010; Aker and Fafchamps, 2014; Tack and Aker, 2014; Jensen and Miller, 2018; Gupta et al., 2019) and enabling value-added services such as mobile money (Jack et al., 2013; Blumenstock et al., 2015, 2018; Suri and Jack, 2016), information provision and reminders (Aker et al., 2012; Cunha et al., 2017; Dammert et al., 2015; Cadena and Schoar, 2011; Dammert et al., 2014; Kast et al., 2013; Karlan et al., 2012, 2016; Zurovac et al., 2011). Yet, research on smartphones remains rare; in a 2019 paper, Hjort and Poulsen (2019) find that connecting Sub-Saharan Africa to high-speed internet cables increased employment, both overall and for the less educated, but the work does not speak to mobile Internet, which is far more common in low-income settings.

In addition, there is also a gap in the literature about who has smartphone access and the

when phone adoption is high, which then has a dampening effect on phone use in the community. Likewise, backlash could be driven by a single naysayer in the community who becomes more likely to discover that women are using phones when penetration is higher.

underlying mechanisms that drive these trends. While the gender gap in mobile phones is well documented, the authors are not aware of any large-scale empirical evidence on how to close these gaps. It is likely that in some developing settings, multiple constraints keep women from accessing and benefiting from the potential returns to mobile phones. For example, in 2022 only 51% of all adult females in India were aware that you can access the internet on a mobile phone, compared to 70% of men. Therefore, the extent to which interventions to improve individual access and ability to navigate mobile technology can overcome usage barriers, improve perceptions around the acceptability of women using phones, and allow women to take advantage of the potential benefits of mobile technology.

Lastly, we contribute to the literature that documents male backlash in response to progressive social shifts. Several recent studies have demonstrated how policy and legal changes can lead to backlash when they threaten existing identities, norms or interests, including Blumenstock et al. (2022); Fouka (2020); Wheaton (2021); Abdelgadir and Fouka (2020); Gottlieb (2016); Mehmood et al. (2022); Bursztyn et al. (2020); Andrew et al. (2022); Brule (2018) and Anukriti et al. (2022). This stands in contrast to evidence of similar changes leading to more progressive views among men or reductions in domestic violence (Beaman et al., 2009; Delaporte and Pino, 2022; Kuipers, 2020). Similarly, several recent papers show that challenges to existing gender norms can subject women to backlash in the form of marital friction, dissolution or violence (Bertrand et al., 2015; Folke and Rickne, 2020; Ashraf et al., 2014; Bobonis, 2011; Angelucci, 2008; Bobonis et al., 2013; Heath, 2014; Erten and Keskin, 2021; Bulte and Lensink, 2019; Luke and Munshi, 2011; Tankard et al., 2019; Chowdhury and Bhuiya, 2004; Tur-Prats, 2021; Tertilt and Van Den Berg, 2012; Anderberg et al., 2016; Baranov et al., 2021). However, the evidence is again mixed (Roy et al., 2019; Angelucci, 2008; Haushofer et al., 2019; Heath et al., 2020; Bobonis et al., 2013; Hidrobo and Fernald, 2013; Hidrobo et al., 2016).

Our results show how training designed to mitigate digital skills gaps could be a useful approach to empowering women in gender-conservative households to utilize phones and take advantage of their potential benefits. Importantly, our findings highlight how investing in women's digital capabilities can yield lasting benefits.

2 Project SKY and Experimental Design

2.1 Project SKY and Smartphone Distribution

In 2017, the Government of Chhattisgarh (GoCG) announced they would distribute free, internet-enabled smartphones to all poor women through a program called Sanchaar Kranti

Yojana (SKY). The aim was to distribute phones to the college students and BPL female head of household in all villages with a population of 1,000 or more as per the 2011 Indian Census. In addition to the smartphone, the GoCG provided a SKY SIM card with 1 GB of free data each month for the first six months. Smaller villages situated in local government units (Gram Panchayats, or GPs) with at least one qualifying village were also eligible. In advance of distribution, the government also commissioned the construction of mobile towers in areas previously outside the footprint of 4G mobile broadband. Approximately 8,000 mobile towers were constructed to ensure all eligible locations had LTE (4G) network coverage.

Distribution occurred between July and October of 2018. Ultimately, SKY distributed smartphones to over 2 million women across 7,720 villages in the state. SKY was well implemented. Program administrative data show that 74% of SKY-eligible villages and just 2% of ineligible villages received that program. The average eligible household received 0.8 phones, close to the program's goal. In our own data, we observe that 61% of women - all of whom were eligible for the program - them reported receiving a phone.

2.2 Sampling

To identify the villages to be included in the study, we use administrative SKY data from Chhattisgarh Infotech Promotion Society (CHiPS) and the Socio-economic Caste Census (SECC).[3] In Raipur district, where our study is based, 377 Gram Panchayats (GPs) and 424 villages were eligible for the SKY program and could be matched in the Census data. Of these, 222 villages meet our village selection criteria and we randomly select 212 to form our final village sample. [4]

To be eligible for inclusion in our study, women had to be: i) under age 40; ii) married; iii) in both the SKY and Census datasets, with at least two family member names listed in the Census (to facilitate identification during surveying) and coded as females in the Census data; iv) belong to households with only one SKY beneficiary. In each of the 212 villages, we sampled 48 eligible women (36 literate and 12 illiterate) for inclusion in the study.[5] From

[3]We subsequently refer to the SECC data as Census data.

[4]First, we matched village data provided by CHiPS to Census village codes and dropped 2 villages which did not match. Then, we dropped 27 villages that we used to pilot our intervention as well as two CHIPS villages where residents could not be matched to Census villages. Thereafter, the sample was restricted to villages with more than 50 and less than 375 SKY beneficiaries which left us with 275 villages. We also dropped 43 villages that did not meet our village-level sampling threshold of 36 literate and 12 illiterate women. We selected this ratio of literate to illiterate women to align with the overall SKY beneficiary ratio for married women under 40. In cases where there was more than one village in a GP, we randomly select one village per GP and we had to drop 10 villages to adhere to this. Using this sampling procedure, 222 villages were identified for inclusion in the sample.

[5]The literacy variable is taken from the 2011 Census data.

among the 48, 8 women (6 literate and 2 illiterate) and their husbands were interviewed both at baseline and at endline, for a total of 1,696 women and 1,696 men ("Sample 1"). An additional 10 women and their husbands (a total of 2,120 women and 2,120 men) were selected to be interviewed only at endline ("Sample 2").[6] So while treatment is assigned among 10,176 women, our survey sample is comprised of 3,816 women and 3,816 husbands.

2.3 Randomization

Before describing the interventions, we outline the randomization. Figure A1 illustrates the design of our intervention which involved both village and individual-level randomization.

Village-level randomization: 32 of 212 villages were randomized to the control group. The remaining 180 out of 212 villages received training on how to use their mobile phones as described in Section 2.4. Half of these villages (90) were randomized to be in *low saturation*: only the 48 women sampled for the study were invited to the training. In *high saturation* villages all sky-eligible women were invited to the training. Therefore, in the 90 villages that were assigned to *high saturation*, a higher proportion of the village was assigned to treatment. On average, 217 women were invited for training in high saturation villages.

The village-level randomization was stratified on three variables: (1) whether the baseline survey was conducted before or after SKY phones were distributed[7], (2) which of the four blocks - the governmental administrative unit below the district - the village was located in and (3) whether the village was above or below median on an index of baseline community perceptions of women's acceptable phone use.[8]

Individual-level randomization: Within each of the 180 training villages, the 48 sampled women are divided into 3 treatments: (1) to only receive phone training (12 respondents); (2) to receive phone training and enroll into the Mor Awaaz service (18 respondents); or (3) to

[6]The remaining 32 women sampled in each village were back-up respondents (20 at baseline, 12 at endline) in case of we were unable to survey women in either Sample 1 or Sample 2. We call these women the "Buffer" sample.

[7]Half of our sample villages were randomized to conduct the baseline survey prior to the distribution of phones, which began in July 2018. The other half was surveyed after phone distribution in their village was completed.

[8]Using our baseline data, we created the index of baseline community attitudes towards women's acceptable phone use using male and female responses to four questions: in your opinion, is it appropriate for [married/unmarried] women to own/use a phone [with/without] supervision?". We generated dummy variables equal to one if the respondent agreed phone use was appropriate, standardized these dummy variables, and then took an average across all questions. Missing values were replaced with village$x$gender means before standardizing and averaging. The final index was the village-level average of both male and female responses.

receive phone training, enroll in Mor Awaaz, and receive an incentive for each Mor Awaaz pull call that is answered by the respondent (18 respondents) answered by the respondent. Both respondents and field staff were blinded to a woman's treatment status when distributing invitations to the phone training camps. Women were offered to enroll into the Mor Awaaz service only once the phone training was completed. And women were informed of the incentive only after they decided whether or not to enroll in Mor Awaaz. The individual-level randomization was stratified on whether the woman the woman was literate or not in the 2011 Census data.

2.4 Interventions

Training: Leveraging the SKY smartphone distribution, our main intervention consists in administering a basic digital literacy training to enhance women's digital capabilities to use smartphones. Within each treatment village, we hosted training camps designed to build the skills women need to engage with smartphones. The 30 minute sessions covered both basic and advanced tasks, such as dialing a number, receiving an incoming call, taking photos, and using Google's voice search feature. For more detail on the content of the training, see Appendix section B.1.

Women were trained in groups of eight to ten and men were not able to attend. In order to keep training groups small and prevent overcrowding, we varied the number of locations where the training could be held, meaning there were between one and three public venues used in each village, depending on the number of invited women.

The training camps occurred between July and October 2019. Of the 10,176 women - "Sample 1", "Sample 2", and "Buffer" - who were sampled for the study, 8640 were assigned to receive training and 4,628 actually attended. An additional 15,235 sky beneficiaries in the 90 high saturation villages were invited for training and 8,711 of them attended the training. So in total, we trained 13,334 women.

Beside the training, we created and administered two additional interventions designed to further promote smartphones ownership and usage among women.

IVR System for Women: The study team developed an IVR service known as "Mor Awaaz", meaning "My Voice" in Chhattisgarhi, in close coordination with the GoCG. Mor Awaaz was implemented among a subset of the women in the treatment villages (i.e., those who received the digital literacy training described above). Enrollment into Mor Awaaz occurred at the end of the training session, with women randomly selected to receive in Mor Awaaz asked to stay longer to be told about and enroll in the service. The Mor Awaaz service had two main components - "pushing," or sending information to women, and "pulling," or

collecting information from them on the delivery of public services that are relevant to them. For a more detailed description of the service, see Appendix section B.2.

The "push" component of the service consisted of recorded voice messages shared with participants on a weekly basis. These voice messages featured characters named *Neetu Didi* and *Bharti Didi*, developed to represent local women who share information about women's health, child nutrition, COVID-19 safety, and government safety net benefits. During the height of the COVID-19 outbreak, we also expanded the content to address vaccine availability, eligibility, and common misconceptions about side effects and efficacy that contribute to vaccine hesitancy. In addition to these generic messages, we also shared customized messages about women's MGNREGS payment statuses with those who recently worked in the public works program for low-income rural households.

For the "pull", or information gathering calls, we set an in-house call center that called women once a month to survey them about their experiences with and perceptions of public service delivery, government programs, and COVID-19. We commonly shared lessons learned from these calls with the Department of Health and Family Welfare and the Department of Panchayat and Rural Development.

Nearly 90% of women who were offered Mor Awaaz at the training camps signed up for the service. In total, 10,329 women were enrolled in Mor Awaaz, 3,048 of whom were in the main sample and the remaining were additional sky-eligible beneficiaries in high saturation villages. The first messages were sent in July 2019 and lasted through the summer of 2022.

Incentives: We offered a randomly selected subset of women in all Mor Awaaz villages small incentives for responding to pull call surveys to better understand the relevance of economic returns to women using the phone. Our prior is that these small financial incentives will amplify the benefits of the service for participants' households, and further legitimize women's use of smartphones. Incentives could be particularly important to help women retain use of the smartphone in her household since other household members may be able to use phones for other income generating opportunities. Before the COVID-19 pandemic, we provided in-kind incentives valued at Rs.100 per pull call survey completed; however, we switched to money orders in early 2020.

2.5 Data

We collected five sets of primary data:

Baseline Data: We conducted the baseline survey from August to November 2018 in Raipur district in Chhattisgarh. The baseline survey collected information from 1,696 women and their husbands ("Core Sample 1") about individual mobile phone usage, actual norms

(first order beliefs) governing women's phone use, perceived norms (second order beliefs), and a module that measures the size of women's social networks. For half of the sample villages, baseline data was collected prior to phone distribution. For the other half, baseline data was collected after phone distribution had been completed in their village.

Training Camp Data: At phone training camps, we collected data on who was invited to the training, who was trained, as well as consent to enroll in Mor Awaaz and the incentives interventions. Prior to the start of the training we also collected data on respondents' health status, mobile phone ownership, first order beliefs around phone use and phone numbers.

Push Call Intervention Data: Between July 2019 and April 2022, we generated 252 push call rounds, spaced out approximately once per week. For each round, we collect administrative data on the pick up rate and duration that a call is listened to.

Pull Call Intervention Data: Throughout the course of the intervention, we generated 10 pull call rounds. In addition to round-specific data, for each round we document whether the call was picked up, the gender of the person who picked up the call, and whether the person picking up the call was our female respondent.

Endline Data: We attempted the endline survey from December 2021 to April 2022 to 3,816 women and their husbands. For both husbands and wives, we collect data on the respodents' phone ownership and use; gender norms and attitudes regarding phones and a broader set of norms about women's role in society; and measures of labor force participation. For female respondents, we additional collect data on their mental health, well-being, gender-based emotional control, empowerment, and social connections and networks.

3 Empirical Approach

3.1 Regression Specification

We estimate two regression equations; in the first, we pool all treatments:

$$y_{iv} = \beta_0 + \beta_1 Treatment_{iv} + Strata_v + \epsilon_{iv} \tag{1}$$

where y_{iv} is the outcome of interest and $Treatment_v$ is a dummy that equals one for villages where women received at least the training intervention (i.e., 180 villages our 212). $Strata_v$ is a vector that includes our randomization strata[9]; ϵ_{iv} an error term.

[9]This includes block, post-SKY distribution, baseline norms, and female literacy. There are four blocks in our dataset: Abhanpur, Arang, Raipur and Tilda. The post-SKY distribution variable is a binary indicator of whether or not the household was surveyed in the baseline after SKY phone distribution. The baseline

The second regression equation allows us to estimate individual-level treatment effects as follows:

$$y_{iv} = \gamma_0 + \gamma_1 TrainingOnly_{iv} + \gamma_2 Training\&MorawaazOnly_{iv} + \gamma_3 Training\&Morawaaz\&Incentive_{iv} + Strata_v + \epsilon_{iv} \quad (2)$$

where y_{iv} is the outcome of interest, $TrainingOnly_{iv}$ refers to women assigned to only received the training; $Training\&MorawaazOnly_{iv}$ is an indicator of whether women were offered to enroll in Mor Awaaz (without any incentives) in addition to receive the training; $Training\ \&\ Morawaaz\&Incentives_{iv}$ refers to whether women were additionally provided with incentives to receive pull calls. So γ_i is the treatment effect of each intervention relative to the control group. $Strata_v$ is again a vector that includes our randomization strata, and ϵ_{iv} an error term. Standard errors are clustered at the village level.

At the bottom of the panel, we test whether the effect of training only (γ_1) is equal to that of training and Mor Awaaz only (γ_2), and whether the effect of training and Mor Awaaz only (γ_2) is indistinguishable from the effect of training, Mor Awaaz, and the incentives treatment arm (γ_3).

The third regression equation allows us to estimate village-level treatment effects as follows:

$$y_{iv} = \alpha_0 + \alpha_1 HighSaturation_{iv} + \alpha_2 LowSaturation_{iv} + Strata_v + \epsilon_{iv} \quad (3)$$

where y_{iv} is the outcome of interest, $HighSaturation_{iv}$ refers to women assigned to high saturation; $LowSaturation_{iv}$ refers to women assigned to low saturation. So α_i is the treatment effect of each intervention relative to the control group. $Strata_v$ is again a vector that includes our randomization strata, and ϵ_{iv} an error term. Standard errors are clustered at the village level.

We construct standardized indices to analyze families of outcomes following Anderson (2008). Specifically, variables are normalized to the mean and standard deviation of the control group. Components are then aggregated, weighed by the inverse covariance of index components. In Appendix C, we explain the definition and sub-components of each index utilized in the analysis.

norms variable is a binary village-level variables that captures whether a village is above median levels for norms related to female phone use. Finally, the female literacy variable is a binary indicator at the household level that captures whether a female respondent is literate. The literacy information was obtained from the Socio Economic and Caste Census (SECC) survey in India.

Summary Statistics and Balance Check: Table A1 shows the randomization balance check for the main individual (Panel A) and household characteristics (Panel B) for the full sample of respondents. As we only collected baseline data for 1696 households, this balance check is based on time-invariant characteristics. Women and men in our sample are respectively 37 and 41 years old on average, these differences being balanced across all treatment groups. The average woman in the control group has 5 years of formal education. Women randomized into the incentives treatment have lower rates of education, but this difference is economically small. Importantly, the joint test fails to reject that there are no important imbalances across all treatments. Men have an average of 7 years of formal education In terms of household characteristics, one in four households belongs to a scheduled caste, and 5% of households belong to a scheduled tribe. The large majority of households belong to other backward castes. All in all, the treatment groups are largely balanced, indicating that the randomization worked well.

In Table A2 and Table A3 we limit the sample to the baseline sample and show a richer set of characteristics for women and their husbands, respectively. In column 1 we show the mean of the control group and in column 2 we show the difference between the control group and all treatments pooled (Specification 1). We observe no economically or statistically significant differences on women's baseline phone use and norms (Table A2). In Table A3, we don't observe differences in men's norms, but treatment men are 11 percentage points more likely to have used a basic phone in the previous 30 days (significant at the 10% level).

Attrition At endline, we attempted to survey 3,816 women and their husbands (a total of 7,632 respondents), of whom 1,696 women and their husbands had been surveyed at baseline. We were able to survey 89.8% of female respondents in the control group. In Table A4, we see that, overall, treatment women were an additional 2.1 percentage points more likely to answer our endline survey, but this difference is not statistically significant at traditional levels. Women assigned to receive just incentives are 2.5 percentage points more likely to be surveyed at endline than women in the control group, significant at the 10% level. But this difference is rather small.

We face a more significant differential attrition problem among men. 78.3% of control males answered the survey. In column 2, we see that treatment men are 5.5 percentage points more likely to be surveyed than control group men (significant at the 1% level). Although attrition is primarily driven by husbands in the training only treatment, there is differential attrition across all treatment groups relative to the control. Differential attrition is, in part, driven by our ability to collect male surveys in a mop-up stage. Surveyors made up to 4 attempts to complete a survey during their visit to a village. If a survey was not completed,

it was re-attempted in a mop-up after the primary push to complete the endline surveys in a particular GP. The survey team was differentially more successful in completing treatment surveys during the mop-up stage. Relative to columns 3 and 4, we recode surveys completed in the mop-up stage to 0 (i.e. as not completed). We see that the differential attrition rate assuming that the mop-up drive was not done is lower for males, although it does not fully explain the difference in completion rates. This differential attrition will not affect the majority of our results since they come from the woman's survey. For male results, we conduct a variety of robustness tests in the paper to argue that differential attrition does not explain results on husbands.

4 Results

4.1 Take Up of Interventions

We analyze the first stage - or the take up of the interventions - in Table A5. As no intervention took place in control villages, there is no chance of non-compliance in the control group. In the table we therefore only analyze take of the treatments by the treated groups. In column 1, the outcome variable is whether a respondent attended the training. 56.6% of those assigned to the training only treatment attended the training. There is no differential take up of training by those assigned to Mor Awaaz or to incentives.

In column 2, we analyze take-up of the Mor Awaaz intervention. The intervention was only offered to respondents who attended the training camp and were assigned to either Mor Awaaz or to incentives. 53.0% of those assigned to Mor Awaaz took up the intervention with no differential take up by the incentives group. Indeed, 95.6% of respondents who attended the camp and were offered Mor Awaaz took it up. Lastly, take up of the incentives intervention was nearly universal, conditional on Mor Awaaz take-up.

Interaction with the Mor Awaaz service is high and stable over time. In Figures A3 and A4 we plot push and pull call pick up rates over the time. We plot participation rates by treatment intervention, limiting the sample to those who signed up for the service. Over the course of the intervention, the average push call pick-up rate is 65% and - with the exception of the COVID lock-down period - this rate remained stable over time (Figure A3). Consistent with the figure, in Table A6 we show that the Mor Awaaz only group and the incentives group pick up at nearly identical rates. But the proportion of calls that is listened to does differ by treatment: the Mor Awaaz group listens to 28% of the messages while the incentive group has a 9.8 percentage point higher listening rate.

For pull calls, surveyors note down the gender of the person who picked up and whether

the person who picked up was the female respondent. As we see in Table A6, on average Mor Awaaz women pick up the call 22% of the time. The incentives intervention raises that rate by 4 percentage points.

Overall, we find that take up of the intervention was high and consistent over time. While incentives do not have a large effect on the extensive margin for either push or pull calls, the incentives group listens to a higher proportion of push calls and exhibits higher rates of survey completion and pick up by the female respondents.

4.2 Impacts on Female Phone Ownership and Usage

We begin by analyzing the impact of our intervention on women's phone ownership and usage. By endline, access to a phone – be it a phone a that our respondent owns or a household phone – is nearly universal in our sample. 97% of women in the control group report having access to a phone and this is not affected by the treatment. The intervention did, however, affect phone ownership. In column 1 (Panel A) of Table 1, we see that treated women score 0.10 standard deviation units higher on a phone ownership index, significant at the 5% level (p-value = 0.019). In Table A7, we show the subcomponents of the phone ownership index and observe that this result is driven by women owning a smartphone. About 15% of women in the control group own a smart phone and the intervention raises that rate by 3 percentage points. Indeed, as shown in column 2, women in the treatment group are almost 50% more likely than women in the control group to own the SKY phone.

The treatment has a large effect on basic and advanced phone usage as well. Basic phone usage includes having used a basic phone, dialed a call, received a call, and sent an SMS. Advanced phone usage includes having used a smartphone, used Whatsapp, took a photo, played a video, searched the web, read information online, sent or received money digitally, or accessed financial information. In columns 3 and 4, we present indexes of phone usage over the past 30 days indexes for these two categories. The intervention led to a 0.15 standard deviation unit increase in both the basic (p-value = 0.000) and also in advanced phone usage in the past month indexes (p-value = 0.002). Consistent with the set of skills that our training intervention was meant to enhance, women who were trained are more likely to have taken a photo, played a video, or searched the web in the previous month (as shown in Table A7). They are also more likely to have sent an SMS or received a call in the previous month - finding that will be discussed in more depth as we present results on the impacts of the intervention on women's social networks.

The pooled effect in Panel A is largely driven by the Mor Awaaz treatment (Table 1; Panel B): women who receive Mor Awaaz score 0.12 standard deviation units higher on

phone ownership. Providing incentives does not increase this treatment effect over and above what is achieved by Mor Awaaz. Although increase in phone ownership is driven by Mor Awaaz and incentives, effects on phone usage (columns 3 and 4 Panel B Table 1) are predominantly concentrated in the training treatment. So the marginal woman who now has her own smartphone is one who would have been utilizing the phone skills she learned in the training regardless.

Men in the control group have significantly higher phone ownership and usage indexes than women (columns 5 - 8). The treatment does not affect their outcomes, meaning that the treatment caused a reduction in the gender ownership and usage gaps. In Figure A5 and Figure A6 we graphically show the shrinking on the gender gap caused by the treatment.

How do effects vary by the proportion of the village that is treated? Women in high saturation enjoy a positive (although not always significant) treatment effect relative to the control. Women in the high saturation villages score on average 0.07 standard deviation units higher on the phone ownership index. While this estimate is large it is not significant at standard levels. They also score 0.11 and 0.09 (both significant at the 1% level) standard deviation units higher on the basic and advanced phone use indexes, respectively. Households in high saturation villages are 3.2 percentage points more likely to still have a SKY phone.

Across the board, women in low saturation consistent show a larger response to the treatment than women in high saturation. Low saturation women score 0.13, 0.19, and 0.22 standard deviation units higher on the phone ownership, basic use, and advanced use indexes than the control group, respectively. These treatment effects are between 63% and 150% larger than those observed in the high saturation treatment. In Section 5, we present evidence consistent with a story of male backlash theory to reconcile these and subsequent differences in treatment effects by saturation.

4.3 Impacts on Social Network

We asked women to report the number of members in her network that belong to the following 3 social groups (i) relatives they do not live with, (ii) friends within their village, and (iii) self-help group (SHG) members.[10] We then asked, for each of these social groups, how many people does she interact with at least once a month and how many she interacts with on the phone at least once a month. In Table 2, we report the effect of the intervention on measures of social interactions.

[10]These three categories are mutually exclusive. We started by asking about relatives that do not live with them. Then, we moved on to friends and ask women to exclude from her friend group count any relatives that do not live with her (as those had already been counted). We asked about SHG group members excluding relatives that do not live with her and friends she has already mentioned.

The average woman in the control group reports having 38.8 relatives that she does not live with, 7.7 friends in her village, and 6.3 SHG member friends. Although the coefficient is positive, we cannot reject at traditional levels that the intervention had no significant impact on an index of the number of friends that women in the treatment reported across the differing social groups (Table 2, column 1). In column 4, we see that women in the treatment group are no more likely to be members of an SHG group. The intervention, therefore, does not expand the size of women's networks. It does, however, affect the intensive margin of interactions.

Treated women engage in more total and over-the-phone interactions with members of their social groups. The intervention led to a 0.08 (p-value = 0.079) and 0.11 (p-value = 0.023) standard deviation unit increase in overall and over-the-phone interactions, respectively (column 3 and 4). Women in the control group interacted with 17.3 social group members (column 7–9 of Table A10) and 5.9 over the phone (column 11–13 of Table A10) in the previous month. Treated women interact with an extra 0.9 relatives overall (significant at the 10% level). This seems to be driven largely by more over-the-phone interactions – they speak to an extra 0.7 relatives on the phone (significant at the 5% level). Results in Panel B of Table 2 indicate that neither the Mor Awaaz or Incentives treatments had additional impacts on social interactions over and above training.

The network interaction results are consistent with what we find in Table 1 with treated women being more likely to pick up a call or send an SMS. Overall, treated women engage with a 1.2 more social group members in a month and 0.9 of these extra interactions are generated over the phone. We conjecture at least two possible channels via which the phone training could have led to a greater number of social connections. The first is via the skills learned at the training. Although the training focused on advanced phone skills, it may also have more generally increased comfort with phone use thereby facilitating the ability of women to connect to people they know using the phone. The second channel is via social interactions generated at the training itself - participating in the training together could have deepened relationships between women and their friends in the same village. As we do not observe how women who attended the training were connected to one another, we cannot distinguish between these channels.

As in the pooled specification, we cannot reject that there is no difference in network size between women in control group and those in either the high or low saturation villages. We do observe differences in treatment effects by saturation, however, when analyzing the indexes of interactions. While the coefficients on total and over-the-phone interactions for high saturation are 0.05 and 0.07, respectively, neither is statistically significant at traditional levels. Women in low saturation, on the other hand, score 0.10 (significant at the 5% level)

and 0.15 (significant at the 1% level) standard deviation units higher on these two indexes. As with the pooled effects, relatively higher phone social interactions in the low saturation group are consistent with the larger treatment effects on phone usage we observe for this group in Table 1.

4.4 Impacts on Female Wellbeing

We collected well-being data along three major dimensions: anxiety and depression, self-esteem, and measures of spousal emotional control. Results are presented in Table 3.

The first column presents an index of anxiety and depression questions which follows the PHQ-4 score designed by Kroenke et al. (2009). Respondents are asked "Over the last two weeks, how often have you been bothered by the following problems?" The problems are (1) Feeling nervous, anxious or on edge, (2) Not being able to stop or control worrying, (3) Feeling down, depressed or hopeless, and (4) Little interest or pleasure in doing things. They answer using a scale of 0 (Not at all) - 3 (Nearly every day). Scores across all 4 questions are added to construct an anxiety and depression index. For consistency of interpretation across columns, we invert the index such that Scores are rated as normal (9-12), mild (6-8), moderate (3-5), and severe (0-2). The average woman in the control group scores 7.7 on the PHQ-4 indicating a mild level of anxiety and depression. 23% of control group women suffer from moderate and 11% from severe anxiety and depression.

The intervention raises the PHQ-4 index by 0.35 points (p-value = 0.006). In Table A11, we show effects on the components of the index. Although across the board the coefficients point to improvements, the results are larger and statistically significant only for reductions in depression.

While we cannot causally link the increase in social interactions we observe to improvements in mental health, psychologists have established that there is positive relationship between the size of a person's social network and level social interactions and their level of depression (see Santini et al. (2015) for a meta analysis).

Turning to the individual treatments, we see that training leads to a 0.51 point increase in the index. Relative to only training, Mor Awaaz reduces the size of that impact by 0.17 points, although this difference is not statistically significant. Nonetheless, women in this group still score 0.34 points higher than women in the control group (p-value = 0.015). The incentives treatment reduces the gain in mental health by another 0.1 points (again, not statistically different from Mor Awaaz). Relative to the control group, incentive women see a gain in mental health of 0.24 points relative to women in the control group We cannot reject that the incentive treatment has no additional effect over Mor Awaaz.

Saturation does not differentially affect the treatment effect on mental health.

The second column of Table 3 presents a self-esteem index based on the Rosenberg Self-Esteem Scale (Rosenberg (1965)). Respondents were asked to state using a 4-point Likert scale how strongly they agree or disagree with the following five statements: (1) On the whole, I am satisfied with myself, (2) At times I think I am no good at all, (3) I feel that I have a number of good qualities, (4) I wish I could have more respect for myself, (5) I take a positive attitude toward myself. Responses were converted such that higher values indicated higher self-esteem and an index that adds the number of points across the questions is presented in column 2. On average, women in the control group score 10.0 points on this scale. We cannot reject that the treatments have no additional impact on women's self-esteem.

In the last column of Table 3, we show an index of emotional control. Respondents were asked on a 3-point Likert scale whether their partner never, sometimes, or always engaged in the following 4 behaviors: (1) He (does/did) not permit you to meet your female friends? (2) He (tries/tried) to limit your contact with your family? (3) He (insists/insisted) on knowing where you (are/were) at all times? (4) He expects you not to contradict him in public? The answers to these questions are combined into a standardized index. We find that none of the treatments affect women's responses about the husbands' level of emotional control.

4.5 Norms

To understand whether the interventions affected gender norms, we ask women their opinions along three dimensions: their own and their spouses beliefs about the appropriateness of phone use for women; their belief about how their communities feel about the appropriateness of women's phone use; and questions related to gender discrimination.

In Table 4, we present women's first order beliefs. Specifically, we ask women whether they think it is or it is not appropriate for women to use phones in the following four scenarios: (1) unmarried woman using phone without supervision, (2) married woman using phone without supervision, (3) unmarried woman using phone with supervision, (4) married woman using phone with supervision. Responses are aggregated into a standardized index presented in column 1 with higher numbers indicating more liberal views.

Nearly twice as many women in the control group believe that supervised phone use is appropriate as unsupervised phone use (Married women: 85.2% to 48.7%; Unmarried women: 72.5% to 38.1%). Whether supervised or unsupervised, women are 10 percentage points more likely to believe that it is appropriate for married women (relative to unmarried women) to use a phone.

The intervention led to a liberalization of women's beliefs about the appropriateness of

phone use. Women who received the intervention have 0.16 standard deviation units more liberal beliefs relative to women in the control group. Disaggregating the questions, we see that this result is driven by greater acceptance of supervised phone use (columns 5 – 7 of Panel A).

In Panel B, we see that Mor Awaaz does not change the size of the treatment effect relative to only training. We do see that women in the incentives treatment score lower than in the other individual treatments, although views are still significantly more liberal than the views of women in the control group. In Panel C, we don't observe important differences in women's own views between low and high saturation.

In Table 5, we show male beliefs about women's phone use. The FOB index is normalized to women in the control group. In column 1, we note that in the control group, men's beliefs about the appropriateness of women's phone use is less liberal than women's. Gaps are especially stark when it comes to unsupervised phone use. The intervention does not change men's first order beliefs. So the treatment served to widen the gender gap in beliefs about women's phone use.

5 Male Backlash

In this section, we aim to understand why we observe larger treatment effects in the villages that were randomized to low saturation compared to high saturation villages. Standard theories of learning or network effects would predict greater phone take-up and engagement the greater the proportion of the village that is treated. Instead, we provide evidence that results can be explained with a model of male backlash: when the a sufficiently large portion of the village is perceived to approve of women's phone use (by attending a publicly visible training, by using phones, etc), men's disapproval rises, reversing the gains made a lower levels of saturation.

In the upper graph of Figure 1, we bin the percentage of village treated into 6 equally-sized bins - the first 3 bins encompass low saturation villages and the other 3 encompass high saturation villages. We estimate the treatment effects on a standardized women's phone use index by proportion of the village that is treated.[11] Consistent with the average results in Panel C of Table 1, phone use rises slightly across the low saturation, and then displays a clear downward sloping trend within the high saturation villages.

These patterns can also be visualized in a more paramaterized regression specification in which phone use is regressed on the fraction treated and the fraction treated squared. These

[11]This index is constructed using basic and advanced phone use and is standardized to the women in the control group.

regressions are estimated separately for the low saturation and the high saturation groups. [12] In the lower graph of Figure 1, we plot the predicted values women's phone use from this parametrized quadratic regression. Note that the gap in estimates over the range of saturation from 0.3-0.42 and 0-0.15 reflects the fact that no villages in either high saturation or low saturation arms fall into these ranges, so we cannot observe what happens over these levels of saturation. The estimated parabola implies that phone use increases sharply from the control group level of 0 up to an apparent tipping point of around 40% treatment, and then declines just as steadily back to zero by the time saturation hits 80%.

In Figure 2, we plot treatment effects on men's (on the left) and women's phone use norms (on the right) by fraction of village treated bin. The graph on the right shows that female norms rise slightly within low saturation villages and then trend downward within the upper median of saturation. As more women in the village are encouraged to adopt, more women deem female phone use acceptable. Meanwhile, male norms over female phone use exhibit a more consistently downward trend almost throughout the distribution of saturation, although not entirely smooth. In Figure 3, we plot the predicted values of men's norms over women's phone use from a quadratic regression on fraction treated. This graph allows us to predict that when saturation is above 70%, not only could norms revert to the control group mean (0), they may even regress to below the counterfactual mean (as seen at a saturation of 80%).

We utilize a second source of variation to analyze backlash: differences in men's baseline acceptance of women's phone use. Specifically, we examine the relationship between husbands' norms and their wives' phone use in low saturation, high saturation, and control households. The results, illustrated in Figure 4, indicate a stronger association between phone use and norms in high saturation areas compared to low saturation areas, supporting the notion of backlash. Greater values on the x axis means more liberal men's views. In high saturation contexts, conservative men become more conservative, while progressive men become more progressive regarding their wives' phone use. Female phone use is lower in high saturation areas relative to low saturation areas among men with below-median levels of baseline acceptance. Conversely, among more progressive men, female phone use is higher in high saturation areas, suggesting that saturation amplifies existing norms of acceptability.

6 Conclusion

This study provides evidence on the effects of training and norms-compatible interventions on women's smartphone adoption and engagement in rural India. The intervention significantly

[12] This graph does not include data from the control group.

increased smartphone ownership and usage among women, enhanced their phone-based interactions with social connections, and improved mental health outcomes, including a reduction in depression. Furthermore, the intervention fostered a liberalization of women's beliefs regarding the appropriateness of female phone use, even though it widened the gender gap in norms, as male attitudes remained largely unchanged. However, the Mor Awaaz service and associated incentives did not yield substantial additional benefits beyond the training itself, underscoring the primacy of skill-building in driving adoption and engagement.

Our analysis also highlights the complex role of community-level dynamics in shaping outcomes. Treatment effects were larger in low saturation villages compared to high saturation ones, suggesting that social norms and male backlash play a significant role in influencing adoption. Men's norms became more conservative and conservative men were less likely to support their wives' phone use as the treated fraction of a village increases, leading to a decline in phone use among women.

References

Abdelgadir, A. and V. Fouka (2020). Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban. *American Political Science Review 114*(3), 707–723.

Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics 2*(3), 46–59.

Aker, J. C. and M. Fafchamps (2014). Mobile phone coverage and producer markets: Evidence from west africa. *World Bank Economic Review 29*(2), 262–292.

Aker, J. C., C. Ksoll, and T. J. Lybbert (2012). Can mobile phones improve learning? evidence from a field experiment in niger. *American Economic Journal: Applied Economics 4*(4), 94–120.

Anderberg, D., H. Rainer, J. Wadsworth, and T. Wilson (2016). Unemployment and Domestic Violence: Theory and Evidence. *Economic Journal 126*(597), 1947–1979.

Anderson, M. L. (2008, 12). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Andrew, A., S. Krutikova, G. Smarrelli, and H. Verma (2022). Gender Norms, Violence and Adolescent Girls' Trajectories: Evidence from a Field Experiment in India. *University of Oxford Department of Economics Working Paper Series*.

Angelucci, M. (2008). Love on the Rocks: Domestic Violence and Alcohol Abuse in Rural Mexico. *The B.E. Journal of Economic Analysis & Policy 8*(1).

Anukriti, S., B. Erten, and P. Mukherjee (2022). Women's Political Representation and Intimate Partner Violence. Technical report.

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.

Baranov, V., L. Cameron, D. Contreras Suarez, and C. Thibout (2021). Theoretical Underpinnings and Meta-analysis of the Effects of Cash Transfers on Intimate Partner Violence in Low- and Middle-Income Countries. *Journal of Development Studies 57*(1), 1–25.

Barboni, G., E. Field, R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2018). A tough call: Understanding barriers to and impacts of women's mobile phone adoption in india.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009). Powerful Women: Does Exposure Reduce Bias. *Quarterly Journal of Economics 4*(124).

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *Quarterly Journal of Economics*, 571–614.

Blumenstock, J., M. Callen, and T. Ghani (2018, October). Why do defaults affect behavior? experimental evidence from afghanistan. *American Economic Review 108*(10), 2868–2901.

Blumenstock, J., O. Dube, and K. Hussain (2022). Can Secular Media Create Religious Backlash? Evidence from Pakistan's Media Liberalization. (February).

Blumenstock, J. E., M. Callen, T. Ghani, and L. Koepke (2015). Promises and pitfalls of mobile money in afghanistan: evidence from a randomized control trial. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*.

Bobonis, G. J. (2011). The Impact of Conditional Cash Transfers on Marriage and Divorce. *Economic Development and Cultural Change 59*(2), 281–312.

Bobonis, G. J., M. González-Brenes, and R. Castro (2013). Public transfers and domesti violence: The role of private information and control. *American Economic Journal: Economic Policy 5*(1), 179–205.

Brule, R. (2018). Reform, Representation and Resistance: The Politics of Property Rights' Enforcement. *Journal of Politics*.

Bulte, E. and R. Lensink (2019). Women's empowerment and domestic abuse: Experimental evidence from Vietnam. *European Economic Review 115*, 172–191.

Bursztyn, L., A. Gonzalez, and D. Yanagizawa-Drott (2020). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review 110*(10).

Cadena, X. and A. Schoar (2011). Remembering to pay? reminders vs. financial incentives for loan payments. Technical Report 17020, National Bureau of Economic Research.

Chowdhury, A. M. R. and A. Bhuiya (2004, 4). The wider impacts of BRAC poverty alleviation programme in Bangladesh. *Journal of International Development 16*(3), 369–386.

Cunha, N., G. Lichand, R. Madeira, and E. Bettinger (2017). What is it about communicating with parents? Technical report.

Dammert, A. C., J. Galdo, and V. Galdo (2015). Integrating mobile phone technologies into labor-market intermediation: a multi-treatment experimental design. *IZA Journal of Labor & Development 4*(1), 1–27.

Dammert, A. C., J. C. Galdo, and V. Galdo (2014). Preventing dengue through mobile phones: evidence from a field experiment in peru. *Journal of Health Economics 35*, 147–161.

Delaporte, M. and F. J. Pino (2022). Female Political Representation and Violence against Women: Evidence from Brazil. Technical report.

Erten, B. and P. Keskin (2021). Trade-Offs? The Impact of WTO Accession on Intimate Partner Violence in Cambodia. *The Review of Economics and Statistics*.

Folke, O. and J. Rickne (2020). All the single ladies: Job promotions and the durability of marriage. *American Economic Journal: Applied Economics 12*(1), 260–287.

Fouka, V. (2020). Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I. *Review of Economic Studies 124*(2), 204–239.

Gottlieb, J. (2016). Why Might Information Exacerbate the Gender Gap in Civic Participation? Evidence from Mali. *World Development 86*, 95–110.

Gupta, A., J. Ponticelli, and A. Tesei (2019). Technology Adoption and Access to Credit Via Mobile Phones. *Available at SSRN*.

Haushofer, J., C. Ringdal, J. P. Shapiro, and X. Y. Wang (2019). Income Changes and Intimate Partner Violence: Evidence from Uncondit. *NBER Working Paper Series 25627*.

Heath, R. (2014). Women's Access to Labor Market Opportunities, Control of Household Resources, and Domestic Violence: Evidence from Bangladesh. *World Development 57*, 32–46.

Heath, R., M. Hidrobo, and S. Roy (2020). Cash transfers , polygamy , and intimate partner violence : Experimental evidence from Mali. *Journal of Development Economics 143*(October 2019), 102410.

Hidrobo, M. and L. Fernald (2013). Cash transfers and domestic violence. *Journal of Health Economics 32*(1), 304–319.

Hidrobo, M., A. Peterman, and L. Heise (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in Northern Ecuador. *American Economic Journal: Applied Economics 8*(3), 284–303.

Hjort, J. and J. Poulsen (2019, March). The arrival of fast internet and employment in africa. *American Economic Review 109*(3), 1032–79.

Jack, W., A. Ray, and T. Suri (2013). Transaction networks: Evidence from mobile money in kenya. *American Economic Review: Papers and Proceedings 103*(3), 356–361.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *Quarterly Journal of Economics 122*, 879–924.

Jensen, R. and N. Miller (2018). Market integration, demand, and the growth of firms: Evidence from a natural experiment in india. *American Economic Review*. Forthcoming.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2016). Getting to the top of mind: How reminders increase saving. *Management Science 62*(12), 3393–3411.

Karlan, D., M. Morten, and J. Zinman (2012). A personal touch: Text messaging for loan repayment. Technical Report 17952, National Bureau of Economic Research.

Kast, F., S. Meier, and D. Pomeranz (2013). Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device. NBER Working Paper No. 18417.

Kroenke, K., R. L. Spitzer, J. B. Williams, and B. Löwe (2009). An ultra-brief screening scale for anxiety and depression: the phq–4. *Psychosomatics 50*(6), 613–621.

Kuipers, N. (2020, 10). The effect of electing female candidates on attitudes toward intimate partner violence. *Journal of Politics 82*(4), 1590–1595.

Luke, N. and K. Munshi (2011, 1). Women as agents of change: Female income and mobility in India. *Journal of Development Economics 94*(1), 1–17.

Mehmood, S., S. Naseer, and D. L. Chen (2022). Why Are Rights Revolutions Rare? (00059).

Rosenberg, M. (1965). Rosenberg self-esteem scale. *Journal of Religion and Health*.

Roy, S., M. Hidrobo, J. Hoddinott, and A. Ahmed (2019). Transfers, behavior change communication, and intimate partner violence: postprogram evidence from rural Bangladesh. *Review of Economics and Statistics 101*(December), 865–877.

Santini, Z. I., A. Koyanagi, S. Tyrovolas, C. Mason, and J. M. Haro (2015). The association between social relationships and depression: A systematic review. *Journal of Affective Disorders 175*, 53–65.

Suri, T. and W. Jack (2016). The long-run poverty and gender impacts of mobile money. *Science 354*(6317), 1288–1292.

Tack, J. and J. C. Aker (2014). Information, mobile telephony, and traders' search behavior in niger. *American Journal of Agricultural Economics 96*(5), 1439–1454.

Tankard, M. E., E. L. Paluck, and D. A. Prentice (2019, 1). The effect of a savings intervention on women's intimate partner violence victimization: Heterogeneous findings from a randomized controlled trial in Colombia. *BMC Women's Health 19*(1).

Tertilt, M. and G. Van Den Berg (2012). Family Violence over the Business Cycle. Technical report.

Tur-Prats, A. (2021, 5). Unemployment and intimate partner violence: A Cultural approach. *Journal of Economic Behavior and Organization 185*, 27–49.

Wheaton, B. (2021). Laws, beliefs and Backlash. *Mimeo, Harvard University*.

Wheaton, B. (2022). Laws, beliefs, and backlash. *Unpublished Working Paper*.

Zurovac, D., R. K. Sudoi, W. S. Akhwale, M. Ndiritu, D. H. Hamer, A. K. Rowe, and R. W. Snow (2011). The effect of mobile phone text message reminders on kenyan health workers adherence to malaria treatment guidelines: a cluster randomised trial. *Lancet 378*(9793), 795–803.

7 Main Figures and Tables

Table 1: Main Phone Ownership & Usage Table

	Female Phone Ownership		Female Phone Usage		Male Phone Ownership		Male Phone Usage		
	Phone Ownership Index	Owns SKY Phone	Basic Use Index (Past 30 days)	Advanced Use Index (Past 30 days)	Phone Ownership Index	Owns SKY Phone	Basic Use Index (Past 30 days)	Advanced Use Index (Past 30 days)	HH has at least one SKY phone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: $\beta_1 Treatment$									
Treatment	0.097**	0.022**	0.150***	0.152***	0.028	0.000	0.010	0.027	0.047***
	(0.041)	(0.009)	(0.041)	(0.049)	(0.047)	(0.011)	(0.045)	(0.073)	(0.017)
Omitted group mean:	-0.000	0.046	-0.000	-0.000	0.696	0.059	0.309	0.254	0.175
SD	(1.000)	(0.209)	(1.000)	(1.000)	(0.997)	(0.237)	(0.945)	(1.324)	(0.380)
N	3457	3457	3500	3499	3457	3457	3168	3169	3450
Panel B: $\gamma_1 T + \gamma_2 T_{CS} + \gamma_3 T_{CS+I}$									
Training only	0.030	0.007	0.111**	0.140**	-0.023	-0.011	0.003	0.020	0.033
	(0.052)	(0.012)	(0.047)	(0.057)	(0.056)	(0.013)	(0.055)	(0.084)	(0.021)
Training & Call Service only	0.120**	0.024**	0.169***	0.133**	0.074	0.011	0.031	-0.018	0.055***
	(0.049)	(0.011)	(0.046)	(0.054)	(0.053)	(0.013)	(0.050)	(0.078)	(0.021)
Training & Call Service & Incentives	0.119**	0.030**	0.156***	0.180***	0.016	-0.003	-0.007	0.074	0.050**
	(0.052)	(0.012)	(0.047)	(0.053)	(0.052)	(0.012)	(0.051)	(0.080)	(0.020)
P-values									
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.030	0.025	0.002	0.010	0.199	0.241	0.814	0.404	0.036
$\gamma_1 = \gamma_2$	0.059	0.088	0.122	0.870	0.044	0.048	0.558	0.536	0.282
$\gamma_2 = \gamma_3$	0.983	0.610	0.728	0.246	0.156	0.162	0.359	0.094	0.803
Panel C: $\alpha_1 HS + \alpha_2 LS$									
High Saturation	0.066	0.012	0.114***	0.087*	0.020	-0.005	0.028	0.016	0.032*
	(0.045)	(0.010)	(0.043)	(0.051)	(0.051)	(0.012)	(0.048)	(0.077)	(0.019)
Low Saturation	0.128***	0.032***	0.186***	0.219***	0.036	0.006	-0.010	0.038	0.063***
	(0.046)	(0.010)	(0.045)	(0.053)	(0.051)	(0.012)	(0.048)	(0.080)	(0.020)
P-values									
$\alpha_1 = \alpha_2 = 0$	0.022	0.005	0.000	0.000	0.767	0.471	0.499	0.873	0.006
$\alpha_1 = \alpha_2$	0.120	0.020	0.031	0.001	0.675	0.220	0.246	0.701	0.056

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Columns (1) - (4) report treatment effects on female phone ownership. Columns (5) - (8) report treatment effects on male phone ownership. Column (9) shows the results for household level SKY phone ownership. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2(Training\&CallService)_i + \beta_3(Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). '

Table 2: Female Social Connectedness

	Network Size Index	Interaction Index	Phone Interaction Index	SHG Membership	Aggregated Travel Index
	(1)	(2)	(3)	(4)	(5)
Panel A: $\beta_1 Treatment$					
Treatment	0.052	0.078*	0.106**	0.018	0.073
	(0.066)	(0.044)	(0.046)	(0.029)	(0.056)
Omitted group mean:	0.000	-0.000	-0.000	0.630	0.000
SD	(1.000)	(1.000)	(1.000)	(0.483)	(1.000)
N	3473	3469	3468	3485	3483
Panel B: $\gamma_1 T + \gamma_2 T_{CS} + \gamma_3 T_{CS+I}$					
Training only	0.070	0.068	0.107*	0.026	0.106
	(0.077)	(0.058)	(0.060)	(0.033)	(0.067)
Training & Call Service only	0.044	0.096*	0.081	0.017	0.044
	(0.070)	(0.054)	(0.053)	(0.031)	(0.062)
Training & Call Service & Incentives	0.048	0.066	0.129**	0.014	0.079
	(0.070)	(0.050)	(0.052)	(0.031)	(0.059)
P-values					
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.842	0.331	0.096	0.890	0.346
$\gamma_1 = \gamma_2$	0.623	0.620	0.654	0.698	0.223
$\gamma_2 = \gamma_3$	0.931	0.519	0.281	0.863	0.403
Panel C: $\alpha_1 HS + \alpha_2 LS$					
High Saturation	0.054	0.052	0.066	0.009	0.048
	(0.070)	(0.050)	(0.052)	(0.032)	(0.062)
Low Saturation	0.050	0.104**	0.146***	0.027	0.097
	(0.069)	(0.050)	(0.052)	(0.031)	(0.061)
P-values					
$\alpha_1 = \alpha_2 = 0$	0.733	0.114	0.019	0.573	0.273
$\alpha_1 = \alpha_2$	0.916	0.280	0.102	0.413	0.351

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Columns (1) to (3) report the treatment effect on social connectedness measured by three social network indices: (i) Network Size index, (ii) Interaction index, (iii) Phone interaction index. Column (4) illustrates the impact on SHG membership/enrolment. Column (5) reports the impact on (independent) aggregated female travel index within and outside of their GP. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Table 3: Main Women Wellbeing

	PHQ-4 Score (Reverse Coded) (1)	Rosenberg Self-Esteem Scale (2)	Emotional Control Index (3)
Panel A: $\beta_1 Treatment$			
Treatment	0.346***	-0.063	-0.005
	(0.123)	(0.105)	(0.069)
Omitted group mean:	7.659	9.927	-0.000
SD	(3.231)	(1.852)	(1.000)
N	3284	3449	3107
Panel B: $\gamma_1 T + \gamma_2 T_{CS} + \gamma_3 T_{CS+I}$			
Training only	0.510***	-0.004	-0.021
	(0.156)	(0.119)	(0.075)
Training & Call Service only	0.343**	-0.070	0.014
	(0.140)	(0.110)	(0.072)
Training & Call Service & Incentives	0.241	-0.096	-0.014
	(0.150)	(0.114)	(0.074)
P-values			
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.011	0.665	0.874
$\gamma_1 = \gamma_2$	0.245	0.440	0.443
$\gamma_2 = \gamma_3$	0.449	0.715	0.513
Panel C: $\alpha_1 HS + \alpha_2 LS$			
High Saturation	0.354**	-0.086	0.033
	(0.139)	(0.109)	(0.073)
Low Saturation	0.338**	-0.040	-0.044
	(0.142)	(0.112)	(0.073)
P-values			
$\alpha_1 = \alpha_2 = 0$	0.021	0.660	0.259
$\alpha_1 = \alpha_2$	0.904	0.512	0.101

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Higher scores indicate better mental health (lower instances of anxiety/depression), better self-esteem, and lower partner emotional control. Column (1) reports the treatment effect on mental health indicators computed using the PHQ-4 instrument. Column (2) reports the effects on self-esteem measured using the Rosenberg Self Esteeem instrument. Column (3) reports the effects on emotional control index components. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Table 4: Female First Order Beliefs (FOB) About Female Phone Use

			Female FOB Unsupervised Use Index Components			Female FOB Supervised Use Index Components	
	Female FOB Index	Female FOB Index (Unsupervised use)	Unmarried No Supervision	Married No Supervision	Female FOB Index (Supervised use)	Unmarried Supervision	Married Supervision
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: $\beta_1 Treatment$							
Treatment	0.155***	0.061	0.037	0.018	0.187***	0.088***	0.045***
	(0.045)	(0.048)	(0.024)	(0.025)	(0.043)	(0.021)	(0.015)
Omitted group mean:	0.000	-0.000	0.381	0.487	-0.000	0.725	0.852
SD	(1.000)	(1.000)	(0.486)	(0.500)	(1.000)	(0.447)	(0.356)
N	3494	3486	3482	3477	3492	3489	3486
Panel B: $\gamma_1 T + \gamma_2 T_{CS} + \gamma_3 T_{CS+I}$							
Training only	0.174***	0.086	0.061**	0.014	0.197***	0.092***	0.048***
	(0.053)	(0.058)	(0.029)	(0.028)	(0.052)	(0.025)	(0.018)
Training & Call Service only	0.190***	0.092*	0.053**	0.028	0.217***	0.104***	0.051***
	(0.051)	(0.054)	(0.026)	(0.027)	(0.048)	(0.023)	(0.017)
Training & Call Service & Incentives	0.108**	0.015	0.004	0.010	0.150***	0.070***	0.036**
	(0.049)	(0.055)	(0.026)	(0.028)	(0.048)	(0.023)	(0.017)
P-values							
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.001	0.165	0.022	0.718	0.000	0.000	0.017
$\gamma_1 = \gamma_2$	0.707	0.892	0.721	0.528	0.640	0.546	0.870
$\gamma_2 = \gamma_3$	0.035	0.076	0.020	0.405	0.074	0.052	0.256
Panel C: $\alpha_1 HS + \alpha_2 LS$							
High Saturation	0.160***	0.047	0.026	0.015	0.206***	0.094***	0.051***
	(0.049)	(0.053)	(0.026)	(0.026)	(0.047)	(0.022)	(0.016)
Low Saturation	0.150***	0.076	0.047*	0.021	0.168***	0.082***	0.038**
	(0.048)	(0.052)	(0.025)	(0.026)	(0.046)	(0.022)	(0.016)
P-values							
$\alpha_1 = \alpha_2 = 0$	0.003	0.342	0.155	0.734	0.000	0.000	0.006
$\alpha_1 = \alpha_2$	0.782	0.467	0.256	0.763	0.275	0.434	0.240

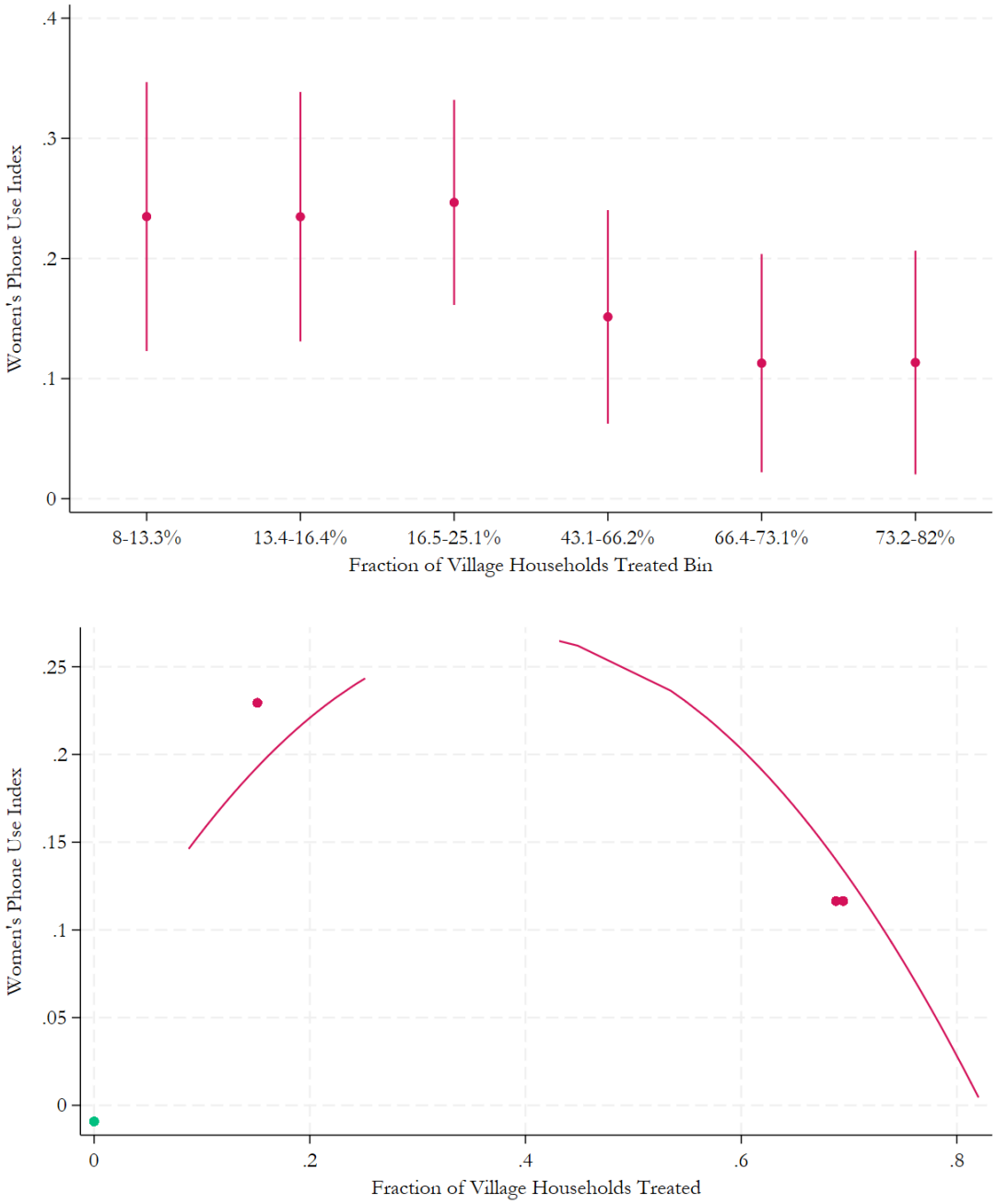
* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.Column (1) reports effects on the female first order beliefs (FOB) index. Columns (2) to (7) report results on sub-indices and components of the aggregated FOB index. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Table 5: Male First Order Beliefs (FOB) About Female Phone Use

			Male FOB Unsupervised Use Index Components			Male FOB Supervised Use Index Components	
	Male FOB Index	Male FOB Index (Unsupervised use)	Unmarried No Supervision	Married No Supervision	Male FOB Index (Supervised use)	Unmarried Supervision	Married Supervision
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: $\beta_1 Treatment$							
Treatment	0.051	0.041	0.017	0.020	0.043	0.010	0.020
	(0.047)	(0.063)	(0.032)	(0.027)	(0.050)	(0.020)	(0.017)
Omitted group mean:	-0.062	-0.152	0.339	0.396	0.030	0.761	0.840
SD	(0.985)	(0.999)	(0.474)	(0.490)	(1.029)	(0.427)	(0.367)
N	3159	3149	3143	3141	3156	3153	3151
Panel B: $\gamma_1 T + \gamma_2 T_{CS} + \gamma_3 T_{CS+I}$							
Training only	0.022	0.028	0.004	0.021	0.009	-0.006	0.011
	(0.057)	(0.073)	(0.036)	(0.032)	(0.058)	(0.024)	(0.020)
Training & Call Service only	0.073	0.021	0.003	0.018	0.087	0.029	0.032*
	(0.053)	(0.068)	(0.034)	(0.029)	(0.055)	(0.023)	(0.020)
Training & Call Service & Incentives	0.049	0.068	0.038	0.022	0.024	0.004	0.013
	(0.053)	(0.069)	(0.034)	(0.030)	(0.054)	(0.022)	(0.019)
P-values							
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.499	0.694	0.390	0.896	0.220	0.267	0.317
$\gamma_1 = \gamma_2$	0.279	0.889	0.951	0.899	0.090	0.084	0.187
$\gamma_2 = \gamma_3$	0.572	0.329	0.132	0.857	0.116	0.132	0.201
Panel C: $\alpha_1 HS + \alpha_2 LS$							
High Saturation	0.029	0.017	0.006	0.009	0.027	0.006	0.014
	(0.051)	(0.066)	(0.033)	(0.028)	(0.054)	(0.022)	(0.019)
Low Saturation	0.073	0.066	0.027	0.031	0.060	0.015	0.026
	(0.051)	(0.067)	(0.033)	(0.029)	(0.052)	(0.022)	(0.019)
P-values							
$\alpha_1 = \alpha_2 = 0$	0.278	0.410	0.502	0.405	0.456	0.735	0.339
$\alpha_1 = \alpha_2$	0.233	0.236	0.285	0.263	0.387	0.550	0.361

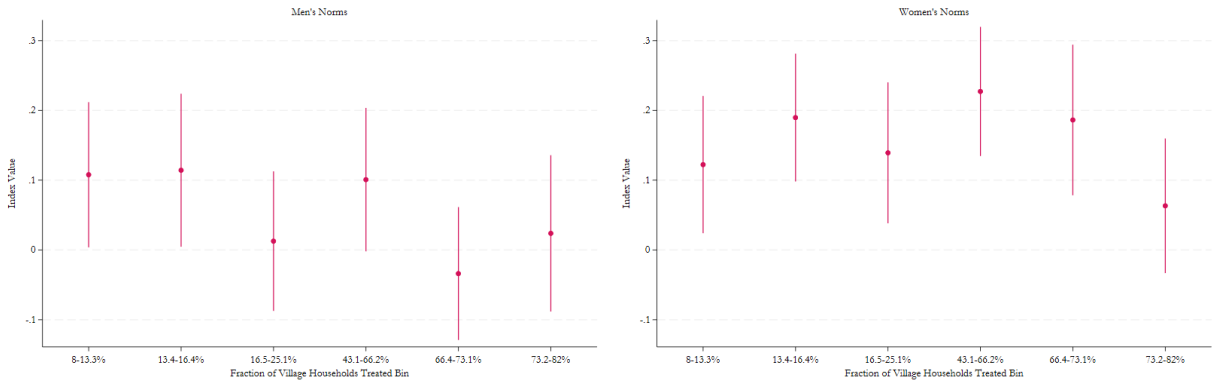
* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.Column (1) reports effects on the male first order beliefs (FOB) index. Columns (2) to (7) report results on sub-indices and components of the aggregated FOB index. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Figure 1: Women's Phone Use and Fraction of Village Households Treated



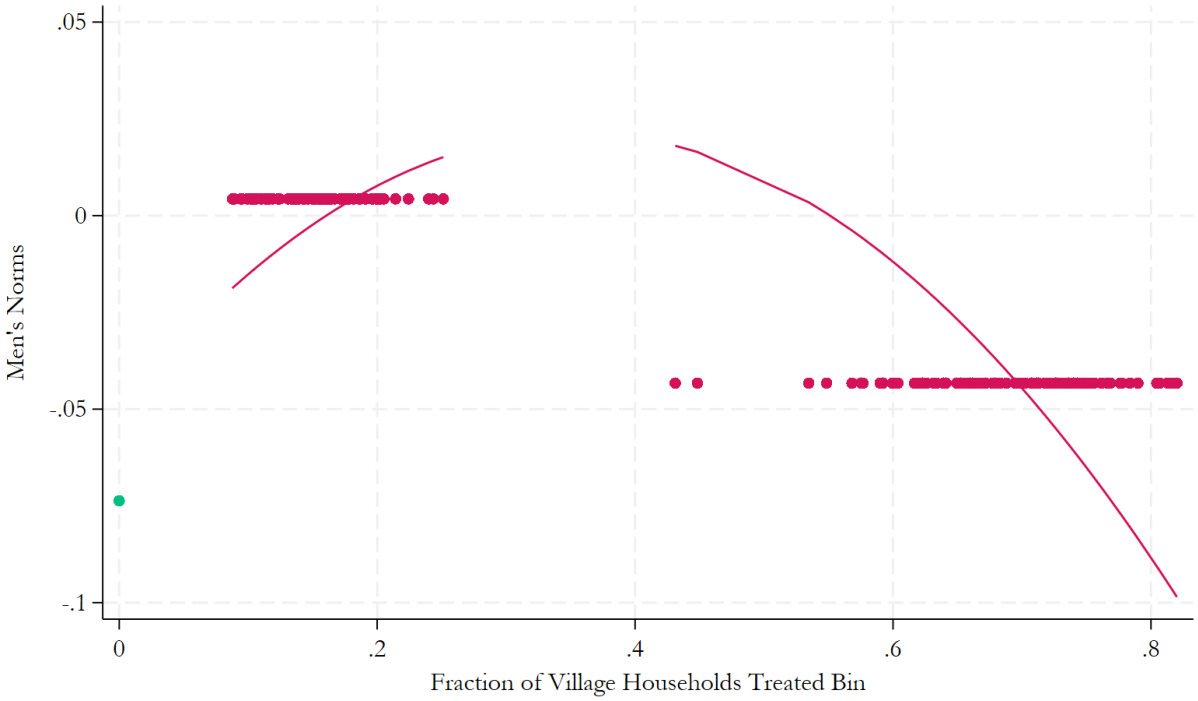
Note: In the upper panel, the omitted group is the control group. Whiskers give 90 percent confidence intervals. The lower panel shows the fitted values of a regression of women's phone use on fraction of village treated and fraction of village treated squared, estimated separately for the low and for the high saturation treatments.

Figure 2: Norms Over Women's Phone Use and Fraction of Village Treated



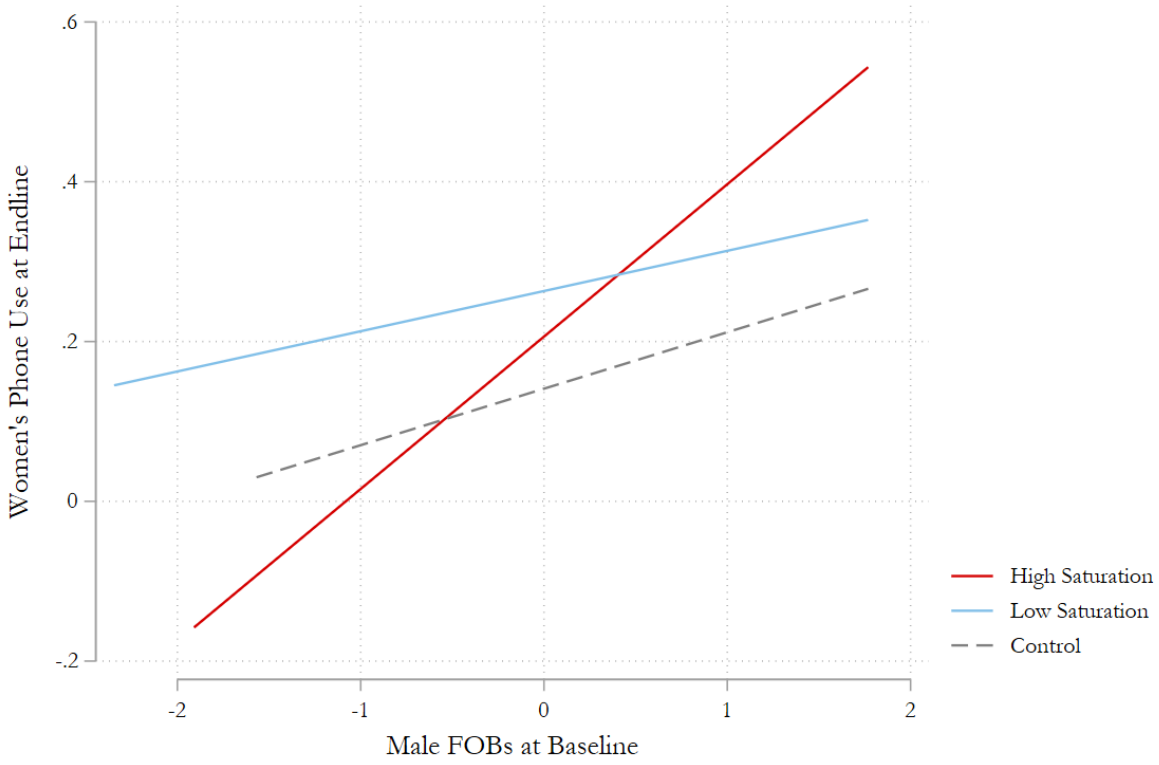
Note: In the upper panel, the omitted group is the control group. Whiskers give 90 percent confidence intervals.

Figure 3: Men's Norms Over Women's Phone Use and Fraction of Village Treated



Note: This graph shows the fitted values of a regression of men's norms over women's phone use on fraction of village treated and fraction of village treated squared, estimated separately for the low and for the high saturation treatments.

Figure 4: Women's Phone Use and Men's Baseline Norms



A Additional Figures and Tables

Figure A1: Randomization Design

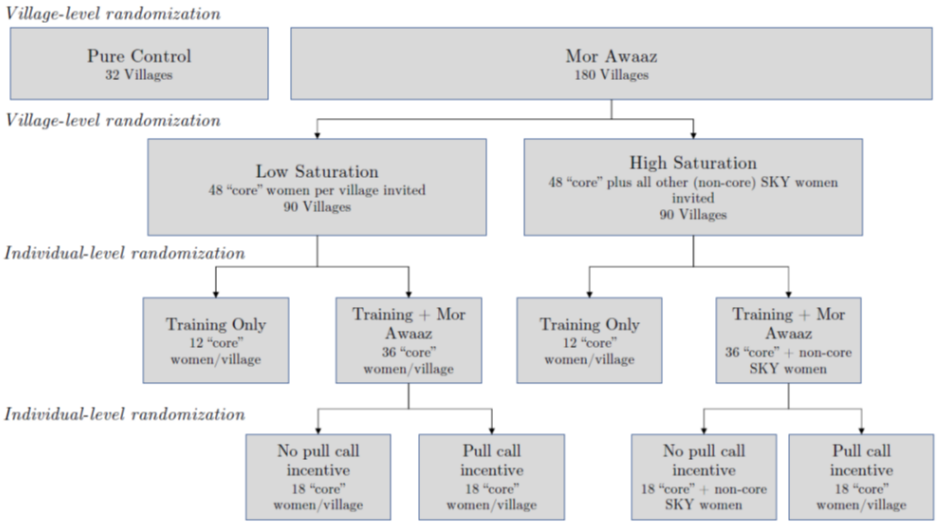


Figure A2: Sampling

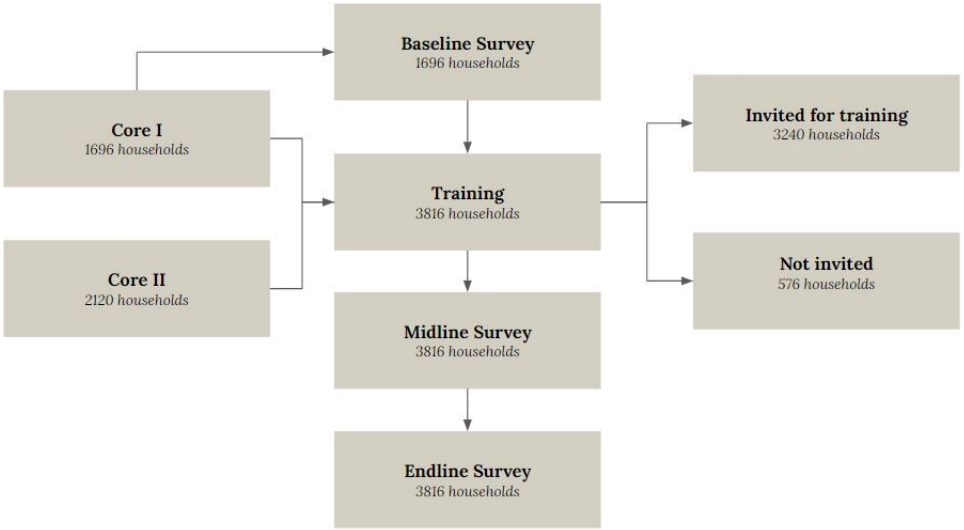


Figure A3: Push Call Pick-up Rate

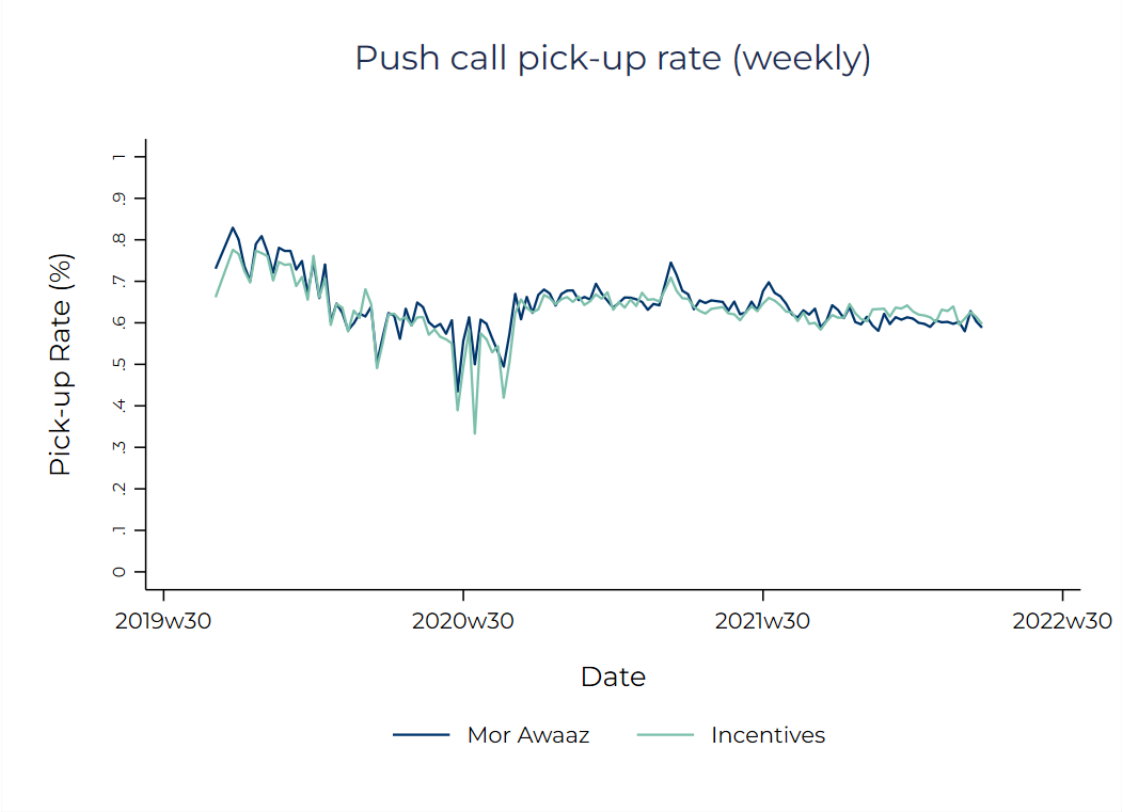


Figure A4: Pull Call Pick-up Rate

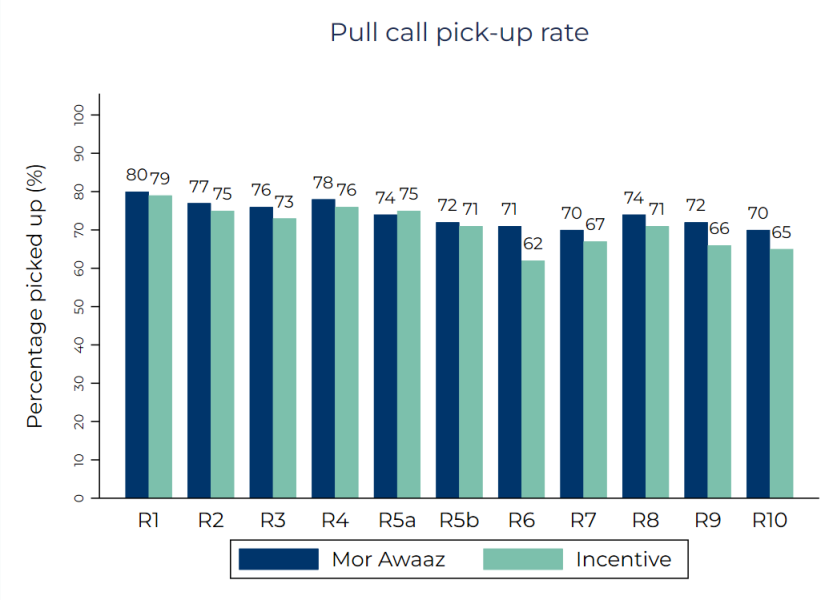


Figure A5: Gender Gap in Phone Ownership

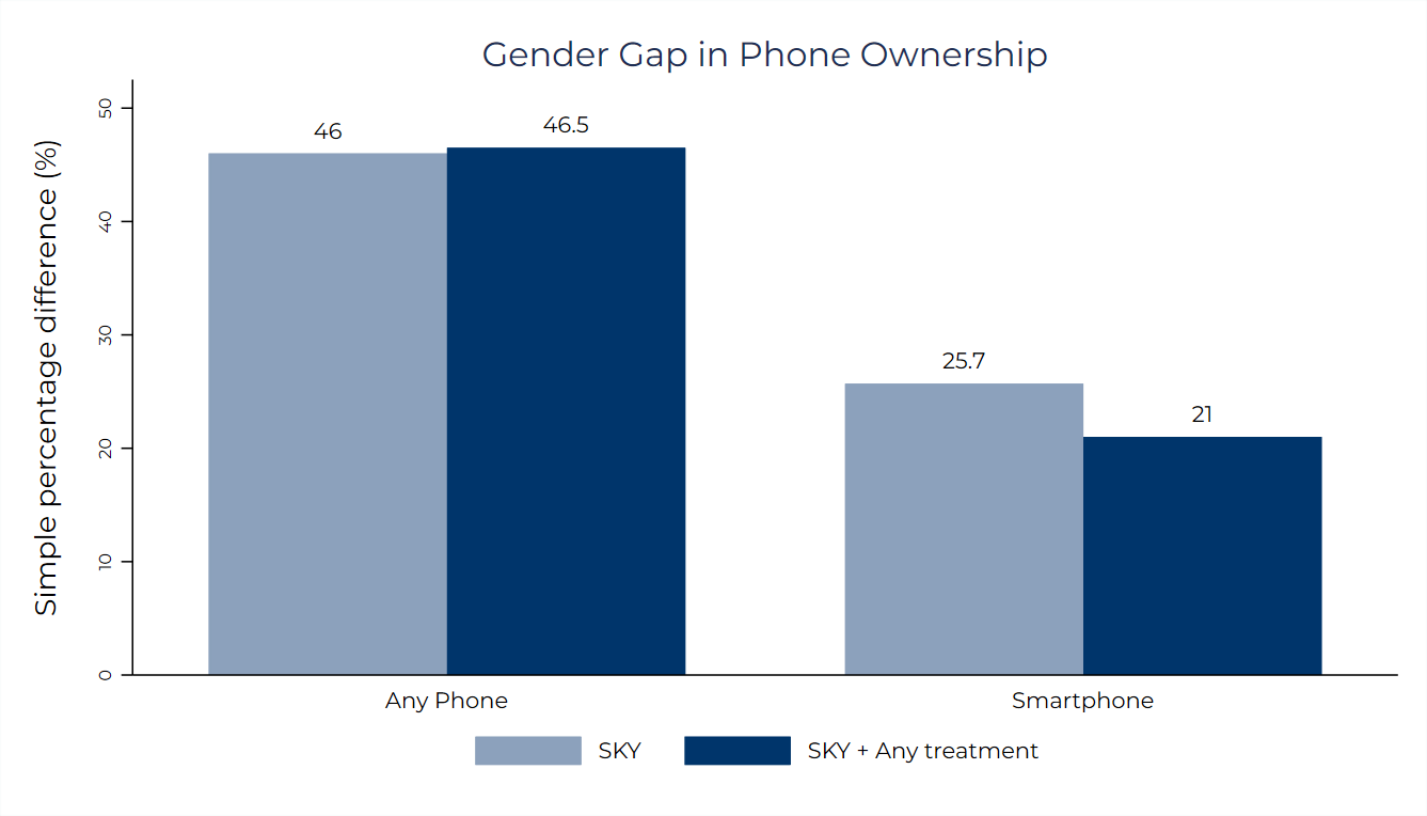


Figure A6: Gender Gap in Phone Usage

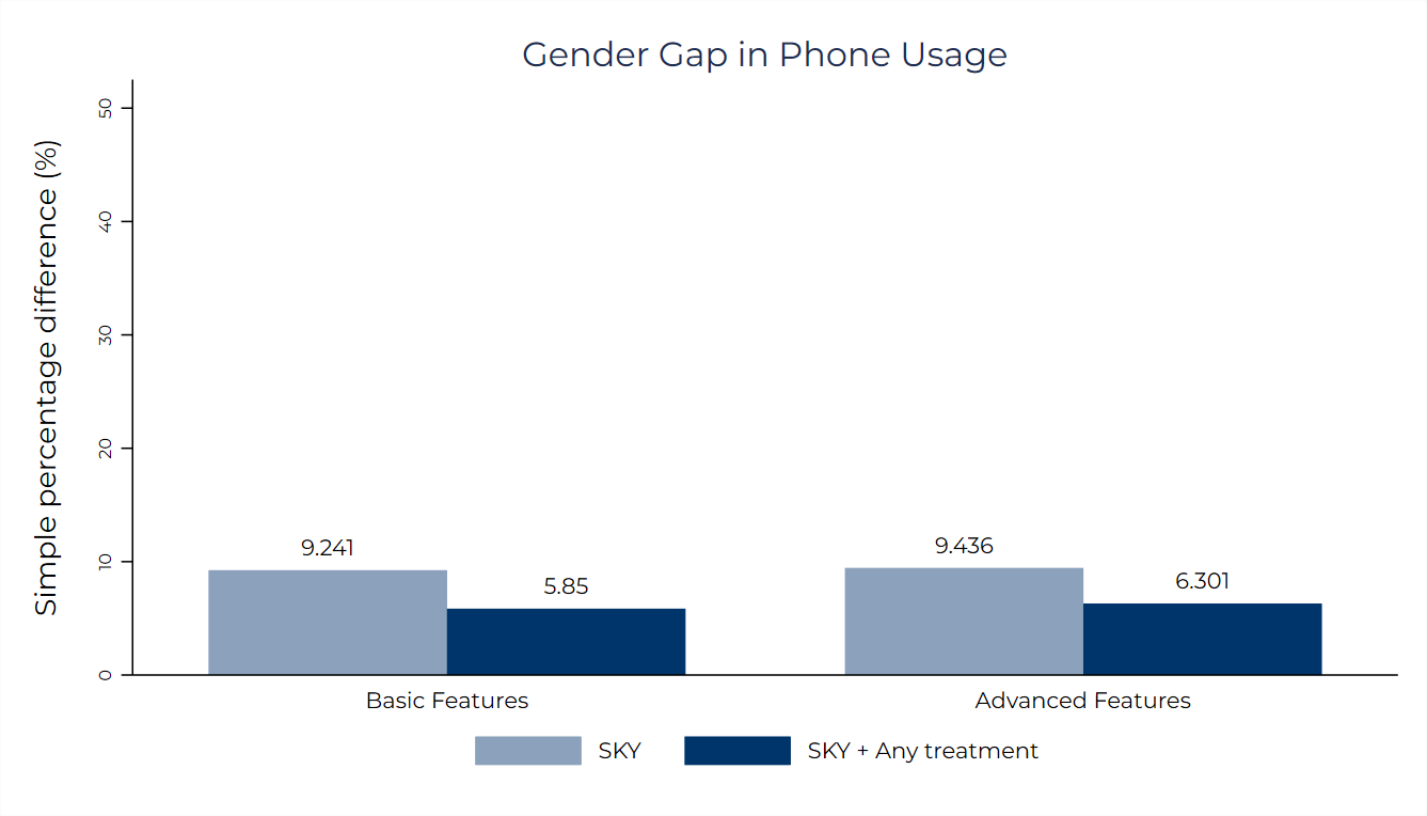


Table A1: Endline Characteristics Balance Table

		Pooled treatment	Individual level randomization				
	Control Mean	Any treatment	Training	Mor Awaaz	Incentives	Joint Test	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Individual Characteristics							
Female: Age	37.075	-0.039	0.001	-0.025	-0.056	0.777	3470
		(0.256)	(0.310)	(0.264)	(0.232)		
Female: Years of education	4.931	0.231	0.299	0.046	-0.273*	0.683	3425
		(0.159)	(0.184)	(0.161)	(0.139)		
Male: Age	41.388	0.215	0.437	-0.325	0.052	0.639	3179
		(0.320)	(0.393)	(0.320)	(0.282)		
Male: Years of education	7.018	0.187	0.283	-0.129	0.001	0.614	3175
		(0.294)	(0.341)	(0.242)	(0.192)		
Panel B: HH Characteristics							
Scheduled Caste	0.207	-0.001	0.000	-0.006	0.010	0.923	3477
		(0.040)	(0.042)	(0.020)	(0.018)		
Scheduled Tribe	0.043	0.007	0.012	-0.009	0.004	0.615	3477
		(0.013)	(0.015)	(0.011)	(0.009)		
Other Backward Caste	0.669	-0.015	-0.006	0.001	-0.027	0.523	3477
		(0.048)	(0.051)	(0.023)	(0.019)		

Table A2: Baseline Characteristics Balance Table - Female

		Pooled treatment	Individual level randomization				
	Control Mean	Any treatment	Training	Mor Awaaz	Incentives	Joint Test	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Phone Engagement							
Have used a mobile phone before	0.883	0.010	-0.001	0.008	0.013	0.431	1696
		(0.024)	(0.028)	(0.019)	(0.018)		
Used basic features before	0.475	-0.005	0.000	-0.011	0.009	0.901	1693
		(0.014)	(0.018)	(0.016)	(0.014)		
Used advanced features before	0.141	0.009	0.010	-0.001	0.001	0.568	1690
		(0.015)	(0.020)	(0.016)	(0.015)		
Used phone in the past 30 days	0.000	0.047	0.031	0.025	-0.006	0.381	1693
		(0.050)	(0.065)	(0.059)	(0.050)		
Used basic features in the past 30 days	0.000	0.039	0.024	0.025	-0.009	0.497	1693
		(0.051)	(0.065)	(0.058)	(0.050)		
Used advanced features in the past 30 days	0.000	0.039	0.052	-0.019	0.001	0.620	1690
		(0.062)	(0.076)	(0.062)	(0.056)		
Panel B: Norms							
Believes female phone use is mostly appropriate	0.000	0.022	0.023	0.013	-0.030	0.927	1694
		(0.062)	(0.074)	(0.060)	(0.056)		
Believes local acceptance of female phone use	0.000	0.084	0.033	0.076	-0.015	0.177	1675
		(0.062)	(0.080)	(0.071)	(0.056)		
Disagrees: Being a good homemaker is women's most important role	0.146	-0.021	0.001	-0.036*	0.014	0.535	1682
		(0.032)	(0.037)	(0.020)	(0.018)		
Disagrees: Men should have the final word about home decisions	0.281	0.025	0.067*	-0.043	-0.028	0.920	1689
		(0.031)	(0.040)	(0.030)	(0.027)		
Disagrees: Men make better political leaders than women	0.476	-0.028	-0.015	-0.004	-0.025	0.231	1668
		(0.033)	(0.041)	(0.034)	(0.029)		
Disagrees: Tertiary education is more important for a boy	0.723	0.017	0.051	-0.064**	0.038	0.535	1679
		(0.038)	(0.041)	(0.028)	(0.027)		
Disagrees: Men have more uses for a phone	0.368	-0.002	0.000	-0.021	0.036	0.694	1679
		(0.033)	(0.041)	(0.034)	(0.030)		

Table A3: Baseline Characteristics Balance Table - Male

		Pooled treatment	Individual level randomization				
	Control Mean	Any treatment	Training	Mor Awaaz	Incentives	Joint Test	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Phone Engagement							
Have used a mobile phone before	0.945	0.016	0.014	-0.007	0.019	0.075	1696
		(0.014)	(0.016)	(0.013)	(0.012)		
Used basic features before	0.641	0.013	-0.006	0.016	0.017	0.064	1696
		(0.013)	(0.017)	(0.016)	(0.014)		
Used advanced features before	0.341	0.020	-0.011	0.035	0.016	0.111	1695
		(0.022)	(0.026)	(0.022)	(0.021)		
Used phone in the past 30 days	0.510	0.119*	0.065	0.026	0.090*	0.008	1696
		(0.062)	(0.073)	(0.057)	(0.052)		
Used basic features in the past 30 days	0.460	0.108*	0.083	-0.012	0.091*	0.010	1696
		(0.056)	(0.068)	(0.058)	(0.049)		
Used advanced features in the past 30 days	0.588	0.080	-0.017	0.086	0.086	0.082	1695
		(0.080)	(0.096)	(0.079)	(0.079)		
Panel B: Norms							
Believes female phone use is mostly appropriate	0.000	0.022	0.023	0.013	-0.030	0.927	1694
		(0.062)	(0.074)	(0.060)	(0.056)		
Believes local acceptance of female phone use	0.000	0.084	0.033	0.076	-0.015	0.177	1675
		(0.062)	(0.080)	(0.071)	(0.056)		
Disagrees: Being a good homemaker is women's most important role	0.071	0.028	0.021	0.032	-0.046**	0.757	1691
		(0.021)	(0.024)	(0.022)	(0.018)		
Disagrees: Men should have the final word about home decisions	0.301	-0.005	-0.004	-0.001	-0.002	0.856	1693
		(0.033)	(0.040)	(0.031)	(0.026)		
Disagrees: Men make better political leaders than women	0.389	-0.023	-0.030	0.030	-0.041	0.327	1671
		(0.038)	(0.045)	(0.032)	(0.028)		
Disagrees: Tertiary education is more important for a boy	0.570	0.055	0.048	0.023	-0.025	0.297	1689
		(0.040)	(0.045)	(0.032)	(0.029)		
Disagrees: Men have more uses for a phone	0.287	-0.026	-0.029	0.014	-0.022	0.252	1678
		(0.028)	(0.035)	(0.029)	(0.027)		

Table A4: Endline Attrition - Mop-ups Recoded (Full Sample)

	No changes to mop-ups		Mop-ups recoded	
	Female Completion (1)	Male Completion (2)	Female Completion (3)	Male Completion (4)
Panel A: $\beta_1 Treatment$				
Treatment	0.021	0.055***	0.017	0.035
	(0.013)	(0.019)	(0.022)	(0.023)
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$				
Training only	0.014	0.072***	0.015	0.047*
	(0.016)	(0.022)	(0.026)	(0.027)
Training & Call Service only	0.021	0.039*	0.002	0.014
	(0.015)	(0.022)	(0.025)	(0.026)
Training & Call Service & Incentives	0.026*	0.060***	0.033	0.049**
	(0.014)	(0.020)	(0.023)	(0.024)
P-values				
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.278	0.006	0.207	0.069
$\gamma_1 = \gamma_2$	0.600	0.038	0.484	0.116
$\gamma_2 = \gamma_3$	0.637	0.165	0.056	0.049

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Columns (1) and (2) show the completion rate for female and male respondents. Columns (3) and (4) are analogous to columns (1) and (2) but we recode the mop-up sample (completion is recoded as 0 if the respondent belongs to the mopup sample). Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 Treatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \gamma_0 + \gamma_1 Training_i + \gamma_2 (Training\&CallService)_i + \gamma_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). Sample: Households from the full endline sample.

Table A5: Training and Mor Awaaz Take-Up

	Attended Training (1)	Enrolled in MA (Camps) (2)
Training & Call Service only	-0.004	
	(0.023)	
Training & Call Service & Incentives	0.011	0.010
	(0.023)	(0.020)
Omitted group mean:	0.566	0.530
Control SD	(0.496)	(0.499)
N	3240	2433

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Standard errors clustered at the village level. Regressions using endline data. All specifications includes strata fixed effects (block, post-Sky distribution, baseline norms and female respondent literacy).

Table A6: Push and Pull Call Engagement

	Push Call Engagement		Pull Call Engagement			
	Call Picked Up	Proportion of script listended	Call Picked Up	Survey Completed	Survey Refused	Respondent Picked Up
	(1)	(2)	(3)	(4)	(5)	(6)
Incentive	-0.006	0.098***	-0.030*	0.077***	-0.105***	0.041***
	(0.015)	(0.013)	(0.015)	(0.015)	(0.015)	(0.013)
Omitted group mean: MA only	0.650	0.282	0.741	0.375	0.484	0.220
SD	(0.477)	(0.395)	(0.438)	(0.484)	(0.500)	(0.414)
N	187604	187592	13879	13879	11020	13879

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Training=1 if they received digital literacy training or moawaaz or incentive (as long as they received dig. lit. training). Mor Awaaz =1 if they received mor awaaz or incentive in addition to digital literacy training. Incentive =1 if they received incentives in addition to dig. lit. training and mor awaaz. Standard errors are clustered at the village level. Regressions are fitted on endline data. The pull call engagement regression specifications include strata fixed effects (block, post-Sky distribution, baseline norms, female literacy) and round fixed effects. The push call engagement regression specifications include strata fixed effects (block, post-Sky distribution, baseline norms, female literacy) and call date fixed effects.

Table A7: Main Phone Ownership Table

	Female Phone Ownership				Male Phone Ownership				
	Phone Ownership Index	Owns Any Phone	Owns Smart Phone	Owns SKY Phone	Phone Ownership Index	Owns Any Phone	Owns Smart Phone	Owns SKY Phone	HH has at least one SKY phone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: $\beta_1 Treatment$									
Treatment	0.097**	0.023	0.030**	0.022**	0.028	0.026	-0.017	0.000	0.047***
	(0.041)	(0.018)	(0.015)	(0.009)	(0.047)	(0.020)	(0.022)	(0.011)	(0.017)
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$									
Training only	0.030	0.006	0.018	0.007	-0.023	0.011	-0.024	-0.011	0.033
	(0.052)	(0.023)	(0.019)	(0.012)	(0.056)	(0.023)	(0.027)	(0.013)	(0.021)
Training & Call Service only	0.120**	0.036*	0.035*	0.024**	0.074	0.039*	-0.014	0.011	0.055***
	(0.049)	(0.021)	(0.018)	(0.011)	(0.053)	(0.023)	(0.025)	(0.013)	(0.021)
Training & Call Service & Incentives	0.119**	0.022	0.034*	0.030**	0.016	0.022	-0.016	-0.003	0.050**
	(0.052)	(0.021)	(0.018)	(0.012)	(0.052)	(0.023)	(0.025)	(0.012)	(0.020)
P-values									
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.030	0.290	0.157	0.025	0.199	0.352	0.843	0.241	0.036
$\gamma_1 = \gamma_2$	0.059	0.136	0.307	0.088	0.044	0.178	0.666	0.048	0.282
$\gamma_2 = \gamma_3$	0.983	0.474	0.971	0.610	0.156	0.372	0.897	0.162	0.803

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. Columns (1) - (4) report treatment effects on female phone ownership. Columns (5) - (8) report treatment effects on male phone ownership. Column (9) shows the results for household level SKY phone ownership. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 Treatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \gamma_0 + \gamma_1 Training_i + \gamma_2 (Training\&CallService)_i + \gamma_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). '

Table A8: Comprehensive Phone Usage Table - Women

		Basic Phone Use Components					Advanced Phone Use Components									Autonomy Index Components	
	Basic Use Index	Used Basic Phone	Dialed a Call	Received a Call	Sent an SMS	Advanced Use Index	Used Smartphone	Used Whatsapp	Took a Photo	Played a Video	Searched the Web	Read Info Online	Send or Received Money	Accessed Financial Info	Phone Use Autonomy Index	Used Phone Outside Home	Does Not Need Permission to Use Phone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Panel A: $\beta_1 Treatment$																	
Treatment	0.150***	0.026	0.030	0.074***	0.019**	0.152***	0.042*	0.022	0.072***	0.069***	0.042**	0.010	-0.000	0.022*	0.090*	0.041	0.019
	(0.041)	(0.021)	(0.020)	(0.017)	(0.008)	(0.049)	(0.022)	(0.017)	(0.018)	(0.020)	(0.019)	(0.015)	(0.008)	(0.012)	(0.050)	(0.027)	(0.019)
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$																	
Training only	0.111**	0.018	0.015	0.075***	0.003	0.140**	0.048*	0.009	0.050**	0.058**	0.031	-0.004	-0.003	0.029*	0.095	0.051*	0.011
	(0.047)	(0.026)	(0.025)	(0.020)	(0.010)	(0.057)	(0.025)	(0.022)	(0.022)	(0.025)	(0.022)	(0.018)	(0.010)	(0.015)	(0.060)	(0.030)	(0.024)
Training & Call Service only	0.169***	0.035	0.038*	0.070***	0.025**	0.133**	0.030	0.028	0.072***	0.075***	0.039*	0.009	-0.003	0.016	0.096*	0.028	0.037*
	(0.046)	(0.024)	(0.023)	(0.018)	(0.010)	(0.054)	(0.024)	(0.020)	(0.021)	(0.023)	(0.021)	(0.017)	(0.009)	(0.014)	(0.053)	(0.028)	(0.021)
Training & Call Service & Incentives	0.156***	0.022	0.032	0.077***	0.022**	0.180***	0.050**	0.024	0.085***	0.071***	0.053**	0.021	0.004	0.022	0.082	0.047	0.007
	(0.047)	(0.024)	(0.023)	(0.019)	(0.010)	(0.053)	(0.023)	(0.020)	(0.021)	(0.023)	(0.021)	(0.017)	(0.009)	(0.014)	(0.055)	(0.029)	(0.022)
P-values																	
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.002	0.545	0.331	0.000	0.021	0.010	0.141	0.446	0.000	0.005	0.088	0.372	0.800	0.222	0.331	0.292	0.188
$\gamma_1 = \gamma_2$	0.122	0.462	0.279	0.735	0.027	0.870	0.320	0.308	0.237	0.455	0.667	0.406	0.982	0.373	0.994	0.273	0.192
$\gamma_2 = \gamma_3$	0.728	0.565	0.722	0.672	0.744	0.246	0.246	0.832	0.468	0.872	0.354	0.397	0.347	0.619	0.729	0.299	0.073

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.Columns (1) - (5) report treatment effects on basic usage. Columns (6) - (14) report treatment effects on advanced usage. Columns (15) - (17) report treatment effects on autonomy to use phone. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 Treatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \gamma_0 + \gamma_1 Training_i + \gamma_2 (Training\&CallService)_i + \gamma_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). ' **??** is the comprehensive male phone usage table. **??** is the comprehensive frequency of female phone usage table. **??** is the comprehensive frequency of female phone usage table. **??** is the comprehensive frequency of male phone usage table.

Table A9: Comprehensive Phone Usage Table - Men

		Basic Phone Use Components					Advanced Phone Use Components							
	Basic Use Index	Used Basic Phone	Dialed a Call	Received a Call	Sent an SMS	Advanced Use Index	Used Smartphone	Used Whatsapp	Took a Photo	Played a Video	Searched the Web	Read Info Online	Send or Received Money	Accessed Financial Info
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: $\beta_1 Treatment$														
Treatment	0.010	0.023	0.008	0.004	-0.013	0.027	-0.011	-0.017	0.014	0.032	-0.024	-0.014	0.009	0.045**
	(0.045)	(0.028)	(0.017)	(0.019)	(0.014)	(0.073)	(0.027)	(0.021)	(0.025)	(0.025)	(0.025)	(0.021)	(0.015)	(0.020)
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$														
Training only	0.003	0.034	-0.005	-0.012	-0.008	0.020	-0.000	-0.028	0.004	0.009	-0.049*	-0.033	0.019	0.047**
	(0.055)	(0.032)	(0.020)	(0.023)	(0.017)	(0.084)	(0.031)	(0.027)	(0.030)	(0.030)	(0.029)	(0.024)	(0.019)	(0.023)
Training & Call Service only	0.031	0.036	0.016	0.014	-0.017	-0.018	-0.035	-0.024	0.035	0.029	-0.026	-0.015	-0.000	0.034
	(0.050)	(0.031)	(0.019)	(0.021)	(0.015)	(0.078)	(0.029)	(0.024)	(0.027)	(0.028)	(0.027)	(0.023)	(0.016)	(0.021)
Training & Call Service & Incentives	-0.007	0.004	0.011	0.005	-0.014	0.074	0.005	-0.002	0.001	0.049*	-0.006	-0.000	0.012	0.055**
	(0.051)	(0.031)	(0.019)	(0.021)	(0.016)	(0.080)	(0.029)	(0.024)	(0.028)	(0.028)	(0.028)	(0.024)	(0.016)	(0.021)
P-values														
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.814	0.413	0.652	0.504	0.680	0.404	0.226	0.510	0.326	0.247	0.199	0.356	0.490	0.074
$\gamma_1 = \gamma_2$	0.558	0.962	0.239	0.133	0.486	0.536	0.131	0.852	0.191	0.418	0.264	0.354	0.196	0.457
$\gamma_2 = \gamma_3$	0.359	0.217	0.729	0.481	0.818	0.094	0.051	0.266	0.104	0.389	0.284	0.419	0.272	0.192

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.Columns (1) - (5) report treatment effects on basic usage. Columns (6) - (14) report treatment effects on advanced usage. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 Treatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \gamma_0 + \gamma_1 Training_i + \gamma_2 (Training\&CallService)_i + \gamma_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). ' **??** is the comprehensive male phone usage table. **??** is the comprehensive frequency of female phone usage table. **??** is the comprehensive frequency of female phone usage table. **??** is the comprehensive frequency of male phone usage table.

Table A10: Comprehensive - Female Social Connectedness

			Network Size Index Components				Interaction Index Components				Phone Index Components		
	Aggregate Networks Index	Network Size Index	Relatives	Friends	SHG Members	Interaction Index	Relatives	Friends	SHG Members	Phone Interaction Index	Relatives	Friends	SHG Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: $\beta_1 Treatment$													
Treatment	0.054	0.052	-1.926	1.001**	0.418	0.078*	0.895*	0.186	0.168	0.106**	0.682**	0.088	0.061
	(0.055)	(0.066)	(3.414)	(0.502)	(0.596)	(0.044)	(0.489)	(0.255)	(0.254)	(0.046)	(0.297)	(0.095)	(0.078)
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$													
Training only	0.068	0.070	-0.111	1.001	0.426	0.068	0.921	0.201	0.133	0.107*	0.660*	0.091	0.073
	(0.069)	(0.077)	(3.811)	(0.647)	(0.649)	(0.058)	(0.622)	(0.329)	(0.289)	(0.060)	(0.365)	(0.115)	(0.088)
Training & Call Service only	0.043	0.044	-2.147	1.022*	0.253	0.096*	1.087*	0.249	0.157	0.081	0.635*	-0.000	0.054
	(0.061)	(0.070)	(3.543)	(0.549)	(0.621)	(0.054)	(0.559)	(0.300)	(0.277)	(0.053)	(0.323)	(0.104)	(0.086)
Training & Call Service & Incentives	0.056	0.048	-2.909	0.980*	0.575	0.066	0.690	0.114	0.201	0.129**	0.743**	0.172	0.059
	(0.060)	(0.070)	(3.628)	(0.537)	(0.634)	(0.050)	(0.550)	(0.275)	(0.279)	(0.052)	(0.336)	(0.106)	(0.083)
P-values													
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.760	0.842	0.660	0.262	0.715	0.331	0.261	0.855	0.914	0.096	0.151	0.122	0.871
$\gamma_1 = \gamma_2$	0.656	0.623	0.383	0.968	0.633	0.620	0.767	0.869	0.917	0.654	0.934	0.319	0.779
$\gamma_2 = \gamma_3$	0.756	0.931	0.691	0.919	0.356	0.519	0.387	0.573	0.817	0.281	0.669	0.026	0.933

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.Columns (1), (2), (6) and (10) reports the treatment effect on the social connectedness, network size, interaction and phone interaction indices. Columns (3) to (5) report effects on the network size index components. Columns (7) to (9) report effects on the interaction index components. Columns (11) to (13) report effects on the phone interaction index components. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 Treatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \gamma_0 + \gamma_1 Training_i + \gamma_2 (Training\&CallService)_i + \gamma_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample). All components are topcoded at the 99th percentile.

Table A11: PHQ-4 Depression and Anxiety Score

			PHQ-4 Anxiety Score Components			PHQ-4 Depression Score Components	
	Total PHQ-4 Score (1)	PHQ-4 Anxiety Score (2)	Feeling nervous, anxious or on edge (3)	Not being able to stop or control worrying (4)	PHQ-4 Depression Score (5)	Feeling down, depressed or hopeless (6)	Little interest or pleassure in doing things (7)
Panel A: $\beta_1 Treatment$							
Treatment	0.346***	0.114	0.064	0.047	0.221***	0.126***	0.093**
	(0.123)	(0.085)	(0.046)	(0.053)	(0.067)	(0.039)	(0.043)
Omitted group mean:	7.659	3.779	1.852	1.929	3.885	1.898	1.990
SD	(3.231)	(1.854)	(1.020)	(1.086)	(1.754)	(1.031)	(1.017)
N	3284	3296	3300	3300	3288	3295	3293
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$							
Training only	0.510***	0.193*	0.070	0.119*	0.306***	0.171***	0.134**
	(0.156)	(0.101)	(0.055)	(0.062)	(0.086)	(0.051)	(0.053)
Training & Call Service only	0.343**	0.116	0.069	0.047	0.218***	0.135***	0.082*
	(0.140)	(0.096)	(0.052)	(0.058)	(0.075)	(0.044)	(0.049)
Training & Call Service & Incentives	0.241	0.060	0.056	-0.000	0.168**	0.088*	0.077
	(0.150)	(0.100)	(0.054)	(0.061)	(0.079)	(0.047)	(0.049)
P-values							
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.011	0.242	0.551	0.103	0.004	0.005	0.093
$\gamma_1 = \gamma_2$	0.245	0.387	0.984	0.151	0.241	0.421	0.270
$\gamma_2 = \gamma_3$	0.449	0.500	0.776	0.321	0.467	0.275	0.885
Panel C: $\alpha_1 HS + \alpha_2 LS$							
High Saturation	0.354**	0.087	0.026	0.057	0.253***	0.140***	0.112**
	(0.139)	(0.092)	(0.050)	(0.056)	(0.075)	(0.043)	(0.048)
Low Saturation	0.338**	0.141	0.103**	0.037	0.189**	0.112**	0.073
	(0.142)	(0.096)	(0.050)	(0.059)	(0.076)	(0.045)	(0.047)
P-values							
$\alpha_1 = \alpha_2 = 0$	0.021	0.340	0.065	0.593	0.003	0.005	0.067
$\alpha_1 = \alpha_2$	0.904	0.491	0.058	0.658	0.352	0.485	0.318

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. We study female mental health using an ultra-brief screening scale for anxiety and depression; the PHQ-4 questionnaire. More information about this questionnaire can be found here. There are a total of 4 items and the scale used is 0 (Not at all) - 3 (Nearly every day). Column (1) reports the effects on the total PHQ-4 score. The total score is determined by adding together the scores of each of the 4 items and range from 0 to 12. We have recoded the options to reverse the valence of the questions. High total score indicates low levels of anxiety/depression. Column (2) reports the effects on the PHQ-4 anxiety score computed by aggregating the first two questionnaire items. Column (5) reports the effects on the PHG-4 depression score computed by aggregating the last two questionnaire items. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Table A12: Rosenberg Self-Esteem Scale

		Rosenberg Self-Esteem Index Components				
	Rosenberg Self-Esteem Scale (RSE) Index	Item 1 Satisfied with self	Item 2* I am no good at all	Item 3 Feel I have good qualities	Item 8* Wish had more respect for self	Item 10 Positive attitude towards self
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $\beta_1 Treatment$						
Treatment	-0.063	0.003	-0.029	-0.049	0.008	0.009
	(0.105)	(0.026)	(0.071)	(0.038)	(0.038)	(0.028)
Omitted group mean:	9.927	2.789	1.367	2.575	0.373	2.816
SD	(1.852)	(0.519)	(1.246)	(0.814)	(0.809)	(0.536)
N	3449	3475	3466	3465	3467	3470
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$						
Training only	-0.004	-0.008	-0.018	-0.003	0.015	0.010
	(0.119)	(0.030)	(0.082)	(0.046)	(0.046)	(0.032)
Training & Call Service only	-0.070	0.002	0.014	-0.073*	0.004	-0.008
	(0.110)	(0.029)	(0.077)	(0.044)	(0.042)	(0.031)
Training & Call Service & Incentives	-0.096	0.010	-0.078	-0.056	0.007	0.024
	(0.114)	(0.029)	(0.077)	(0.043)	(0.042)	(0.030)
P-values						
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.665	0.900	0.346	0.227	0.987	0.530
$\gamma_1 = \gamma_2$	0.440	0.695	0.597	0.097	0.759	0.473
$\gamma_2 = \gamma_3$	0.715	0.731	0.080	0.622	0.909	0.148
Panel C: $\alpha_1 HS + \alpha_2 LS$						
High Saturation	-0.086	-0.006	-0.057	-0.040	0.017	0.003
	(0.109)	(0.028)	(0.076)	(0.041)	(0.041)	(0.030)
Low Saturation	-0.040	0.012	0.000	-0.058	-0.002	0.014
	(0.112)	(0.027)	(0.076)	(0.043)	(0.042)	(0.030)
P-values						
$\alpha_1 = \alpha_2 = 0$	0.660	0.601	0.529	0.393	0.830	0.800
$\alpha_1 = \alpha_2$	0.512	0.317	0.293	0.572	0.571	0.559

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01.We use the Rosenberg Self-Esteem Scale (RSE) to measure female self-esteem. The full RSE inventory has 10 items answered in a four point likert scale ranging from strongly agree to strongly disagree. We use a subset (items 1,2,3,8 and 10) of this inventory for our study. Item that are marked with asterisk * are reversed in valence in the original questionnair and are recoded accordingly (There are 5 questions and the scale used is 0-3, so the total score will range from 0-15). Low total score indicates low levels of self esteem. Column (1) reports the effects on the total Rosenberg SES score. Columns (2) to (6) report effects on the individual items of the RSE inventory. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2(Training\&CallService)_i + \beta_3(Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

Table A13: Emotional Control Index

		Emotional Control Index Components			
	Emotional Control Index	Not permitted to meet friends	Contact with family is limited	Insisted to report whereabouts at all times	Expected to not contradict in public
	(1)	(2)	(3)	(4)	(5)
Panel A: $\beta_1 Treatment$					
Treatment	-0.005	0.009	-0.025	-0.058	0.066
	(0.069)	(0.045)	(0.043)	(0.041)	(0.048)
Omitted group mean:	-0.000	0.372	0.347	0.434	0.651
SD	(1.000)	(0.639)	(0.640)	(0.667)	(0.786)
N	3107	3103	3106	3101	3100
Panel B: $\gamma_1 T + \gamma_2 T + CS + \gamma_3 T + CS + I$					
Training only	-0.021	0.013	-0.034	-0.019	0.065
	(0.075)	(0.049)	(0.048)	(0.047)	(0.056)
Training & Call Service only	0.014	0.004	-0.026	-0.086**	0.062
	(0.072)	(0.047)	(0.046)	(0.044)	(0.051)
Training & Call Service & Incentives	-0.014	0.012	-0.018	-0.057	0.070
	(0.074)	(0.048)	(0.046)	(0.044)	(0.052)
P-values					
$\gamma_1 = \gamma_2 = \gamma_3 = 0$	0.874	0.984	0.903	0.085	0.588
$\gamma_1 = \gamma_2$	0.443	0.782	0.792	0.035	0.930
$\gamma_2 = \gamma_3$	0.513	0.747	0.766	0.279	0.806
Panel C: $\alpha_1 HS + \alpha_2 LS$					
High Saturation	0.033	-0.004	-0.061	-0.072	0.053
	(0.073)	(0.048)	(0.046)	(0.043)	(0.051)
Low Saturation	-0.044	0.023	0.012	-0.044	0.078
	(0.073)	(0.048)	(0.046)	(0.044)	(0.052)
P-values					
$\alpha_1 = \alpha_2 = 0$	0.259	0.734	0.046	0.237	0.317
$\alpha_1 = \alpha_2$	0.101	0.452	0.016	0.344	0.503

* p≤ 0.10, ** p≤ 0.05, *** p≤ 0.01. There are 4 questions and answered in a 3 point likert scale ranging from 0 (Never) - 2 (Always). Therefore, low aggregated scores indicate low instances of emotional control. Column (1) reports the effects on the aggregated emotional control index. Columns (2) to (5) report the effects on emotional control index components. Abbreviations: T = training, CS = call service (Mor Awaaz), I = incentives, HS = high saturation village, LS = low saturation village, FT = fraction treated. Regression specification: (i) Panel A: $Y_i = \beta_0 + \beta_1 AnyTreatment_i + \gamma_s + \epsilon_i$, (ii) Panel B: $Y_i = \beta_0 + \beta_1 TrainingOnly_i + \beta_2 (Training\&CallService)_i + \beta_3 (Training\&CallService\&Incentive)_i + \gamma_s + \epsilon_i$, (iii) Panel C: $Y_i = \alpha_0 + \alpha_1 HighSaturation_i + \alpha_2 LowSaturation + gamma_s + \epsilon_i$. γ_s is the strata (block, post-SKY distribution, baseline norms, SECC female literacy) fixed effects. Standard errors are clustered at the village level. Regressions fitted on endline data (final sample).

B Appendix

B.1 Digital Literacy Training

Recruitment: The day before the digital literacy training occurred, women were invited to attend in-person by professional enumerators. Generally, enumerators attempted to find women while they were at home, but they were also able to approach women at work or another location if it was within the borders of the treatment village. Enumerators made up to three attempts to invite each woman; during the invitation process, enumerators would explain the purpose of the training as well as leave an invitation card to remind them of the date, time, and location of the training (see Figure A7 for an example of the front of the invitation card). In cases in which the woman could not be contacted after three attempts, the invitation card was left with one of their household members. In total, 23,875 women were targeted for training; 21,328 women, or 89%, were invited, and 13,334 women, or 62.5% ultimately attended a training.

Figure A7: Sample Invitation Card Front



Training Materials Development: The content of the training was developed through a process of qualitative interviews, testing material during pilot training sessions, and review by expert partners including the United States Agency for International Development (USAID) and Internet Saathi, a joint initiative of Google India, Intel and Tata Trusts that focuses on providing digital training to rural women in India (see Figure A8 for a sample of what the training materials looked like). Professional graphic designers were employed to develop the branding of the materials. The training sessions occurred after the distribution of the SKY phones, so women were instructed to bring their devices with them. If they were unable to bring the SKY phone specifically, they were encouraged to bring whatever phone they used most often. Women were also provided a small incentive to attend - a tiffin box.

Training Content and Implementation: Training material was printed out in large format for women to see, was explained orally, and was demonstrated on a device. Women performed the actions in the training directly on their own phones, while trainers were available to assist them. Women were also encouraged to interact with each other, and test out the features they were learning, like taking photos or making/receiving calls. Training lasted approximately 30 minutes, and were limited to groups of 10 to ensure there was enough capacity for trainers to assist as needed. The training materials, in Chhattisgarhi, were also provided in a pamphlet for women to take home. The training focused on two broad categories of tasks: basic uses (turning a phone on and off, unlocking a phone, dialling a number, receiving an incoming call) and advanced uses (taking and viewing photos, and using Google's voice search feature). (see Figure A9 for a photo of what a typical training session looked like).

Figure A8: Sample Training Material



Figure A9: Photo from Training Session



B.2 Mor Awaaz Service

Recruitment: Women who were randomly assigned to receive the Mor Awaaz service were recruited directly after the conclusion of the training camp. They were provided with information about the service both orally from trained enumerators and were provided with a written brochure about the service. Women had the option to opt-in or opt-out of the service, and could choose which mobile phone number to be contacted on. The service was provided for free. The services consisted of two types of calls: push calls and pull calls. Both types of services began in mid-2019 and ended in mid-2022.

Push Calls The push calls were weekly, pre-recorded messages sent to womes' phones that covered a variety of topics. They last anywhere between one and three minutes; and were professionally recorded. These voice messages featured characters - Neetu Didi and Bharti Didi - and were centered on three types of information: nutrition/health, COVID-19 response, and government services. The exact messages women would receive was also dependent on what life stage they were in, as indicated during sign-up. Women were categorized into one of the following categories: women with children under 6, women with daughters aged 10 - 18, Breastfeeding women, Pregnant women, Other (none are true).

- Health and nutrition messages were developed in collaboration with the Ministry of Women and Child Development in Chhattisgarh. Topics included anemia, family

planning, micro- and macronutrients, menstruation, etc.

- COVID-19 messages include information about how it spread, how to get vaccines, information on lockdowns, etc

- Government services messages included information on how to sign-up for services, including during COVID-19. There was also a set of women who participated in the rural workfare program called Mahatma Gandhi National Rural Employment Guarantee Act 2005 (MGNREGS), who received personalized messages on when their payments might be expected. We did this by scraping the publicly available MGNREGS database for those who shared their job card (program registration) information with us at the time of enrollment, and translating this to a voice message in the voice of Bharti Didi.

Messages were sent using a professional service, which also allowed us to track information on how participants engaged with the service. This includes pick-up rates, the gender of the individual who picked up the phone and the duration of the call listened to.

Pull Calls Pull calls were delivered by professional enumerators approximately once a month to survey them about their experiences with and perceptions of public service delivery, government programs, and COVID-19. The content of these surveys were also developed in conjunction with the Department of Health and Family Welfare and the Department of Panchayat and Rural Development. We shared aggregated statistics from our pull calls with these departments to help them better provide and advertise their services. Pull calls generally lasted approximately 15 minutes.

A randomly selected subset of women in all Mor Awaaz villages received small incentives for responding to pull call surveys to better understand the relevance of economic returns to women using the phone. Before the COVID-19 pandemic, in-kind groceries were delivered in-person. However, once this became impossible due to lock-downs, monetary incentives were delivered by the India Post to deliver money orders to respondents. The incentives were implemented effectively, with a delivery rate of over 98%.

C Variable and Index Definitions

C.1 Phone Engagement

C.1.1 Construction of the standardized phone engagement indices

1. **Phone Ownership Index**: This aggregated index contains three components: (i) owns any phone - information extracted from the phone roster and constructed using information from ownership and possession questions, (ii) owns smart phone, and (iii) owns SKY phone.

2. **Use Index (Last Month)**: This composite index captures basic and advanced phone usage in the past month. It contains components from two sub-indices: (i) basic use index and (ii) advanced use index. When constructing the composite index, we aggregate the individual components using GLS weighting instead of aggregating the sub-indices.

   (a) **Basic phone use index (Last month)**: This index captures basic phone usage in the past month (at the time the respondent is surveyed). The components of this index includes: (i) used button phone, (ii) dialed a call, (iii) received a call, and (iv) sent an SMS.

   (b) **Advanced phone use index (Last month)**: This index has eight components that aim to capture advanced phone uses in the past month (at the time of the survey). The components of this index are: (i) used smart phone, (ii) used WhatsApp, (iii) took a photo, (iv) played a video, (v) searched the web, (vi) read information online, (vii) sent or received money on the phone, and (viii) accessed financial information on the phone.

C.2 Social Network and Connectedness

C.2.1 Construction of standardized social network indices

1. **Aggregate Networks Index**: This composite index captures female social network and connectedness. It contains components from three sub-indices, (i) network size index and (ii) interaction (in-person) index, and (iii) phone interaction index. When constructing the composite index, we aggregate the individual components using GLS weighting instead of aggregating the sub-indices.

2. We consider three types of social network groups for these indices: (i) relatives that

respondents do not live with, (ii) close friends within their village and (iii) self-help group (SHG) members.

3. **Network Size Index**: This index captures the size of female network or social circle. We ask female respondents "How many people do you know who are your [social network group]?" We ask this question for all three groups (relatives, friends, and shg members) and aggregate them to form the network size index.

4. **Interaction Index**: This index captures the in-person interaction female respondents have with their social circle. We ask female respondents "Of these [number of people they know] people, how many do you speak with at least once a month?"

5. **Phone Interaction Index**: This index captures the phone interaction female respondents have with their social circle. We ask female respondents "Of these [number of people they know] people, how many do you speak with on the phone at least once a month?"

C.3 Norms

C.3.1 Construction of standardized norms indices

1. We want to study if our interventions shift social norms surrounding female phone use. We consider four scenarios to measure first order (what they think about something) and second order (what they think others think about something) beliefs on female phone use: (i) unmarried woman using phone without supervision, (ii) married woman using phone without supervision, (iii) unmarried woman using phone with supervision, (iv) married woman using phone with supervision. We ask the respondents if these four scenarios are appropriate (asked separately).

2. **FOB index**: We construct a FOB (first order belief) index for both women and men. This composite index aims to capture the respondent's own belief about women's phone use under different circumstances (marital status and supervision).

C.4 Female Wellbeing and Mental Health

C.4.1 PHQ-4: The Four-Item Patient Health Questionnaire for Anxiety and Depression

1. Questions [Over the last two weeks, how often have you been bothered by the following problems?]:

- Item 1: Feeling nervous, anxious or on edge
- Item 2: Not being able to stop or control worrying
- Item 3: Feeling down, depressed or hopeless
- Item 4: Little interest or pleasure in doing things

2. The scale used is 0 (Not at all) - 3 (Nearly every day)

3. Total score $\geq$3 for first 2 questions suggests anxiety. Total score $\geq$3 for last 2 questions suggests depression.

4. Total score is determined by adding together the scores of each of the 4 items. Scores are rated as normal (0-2), mild (3-5), moderate (6-8), and severe (9-12). Low total score indicates low levels of anxiety/depression

5. *Reference: Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics. 2009;50(6):613-21. From Principles of Neuropathic Pain Assessment and Management, November 2011.*

C.4.2 Rosenberg Self Esteem Scale

1. A 10-item scale that measures global self-worth by measuring both positive and negative feelings about the self. The scale is believed to be uni-dimensional. All items are answered using a 4-point Likert scale format ranging from strongly agree to strongly disagree.

2. We use the following five items from the Rosenberg Self-Esteem Scale Questionnaire:

   - Item 1: On the whole, I am satisfied with myself.
   - Item 2: At times I think I am no good at all. **
   - Item 3: I feel that I have a number of good qualities.
   - Item 8: I wish I could have more respect for myself. **
   - Item 10: I take a positive attitude toward myself.

3. The scale we use is a 4-point Likert scale format ranging from strongly agree to strongly disagree. Some items are reverse in valence (denoted with **).

   - For items 1,3 and 10: 0 (Strongly Disagree) - 3 (Strongly Agree)
   - For items 2 and 8 that are reverse in valence: 0 (Strongly Agree) - 3 (Strongly Disagree)

4. The total score will range from 0-15. Low total score indicates low levels of self esteem.

5. *Reference: Rosenberg, M. (1979). Conceiving the Self. New York: Basic Books.*

C.4.3 Emotional Control Questions

1. We ask the following four questions related to emotional control:

    - Item 1: He (does/did) not permit you to meet your female friends?
    - Item 2: He (tries/tried) to limit your contact with your family?
    - Item 3: He (insists/insisted) on knowing where you (are/were) at all times?
    - Item 4: He expects you not to contradict him in public?

2. We recode the responses into a 3 point likert scale ranging from 0 (Never) - 2 (Always). Therefore, low aggregated scores indicate low instances of emotional control.

3. We use GLS weighted indices when aggregating the components. Missing values are replaced with 0 (control subsample mean; control females)