Pre-analysis Plan:

Gender Competition and Norms Around Women's Work

Charity Troyer Moore
Suanna Oh
Rohini Pande
Simone Schaner

November 2024

1 Motivation

We aim to understand how expressed gender norms - particularly around whether it is appropriate for women to work for pay - are influenced in the short term by private economic incentives. Working in socially conservative communities of rural India, this study will assess how gender norms evolve when men and women from the same household have to compete for valuable digital jobs new to communities.

Our study offers mobile app-based gig work, addressing many issues that typically hinder women from working - e.g., concerns about privacy, flexibility, and mobility - but are also attractive to men due to high remuneration. We cross-randomize gender reservations for jobs (male, female, or open to both genders) and high and low-wage opportunities across community clusters.

We lay out several theories of norm change against which we benchmark our findings, aiming to distinguish between different drivers of change. These theories include strategic norm enforcement, which posits that when men have to compete with women for jobs, they will express more conservative views related to women's work, especially if the jobs are economically attractive; messaging inference, which posits that reservation and wage policies that signal a job is "appropriate" for a group will lead to norm updating; and cognitive dissonance, which suggests that conservative norms will liberalize when women have an incentive to take up jobs. We will also examine consequences of gender reservations for women's actual job take-up and productivity, and preferences over reservation policies for future potential work in the community.

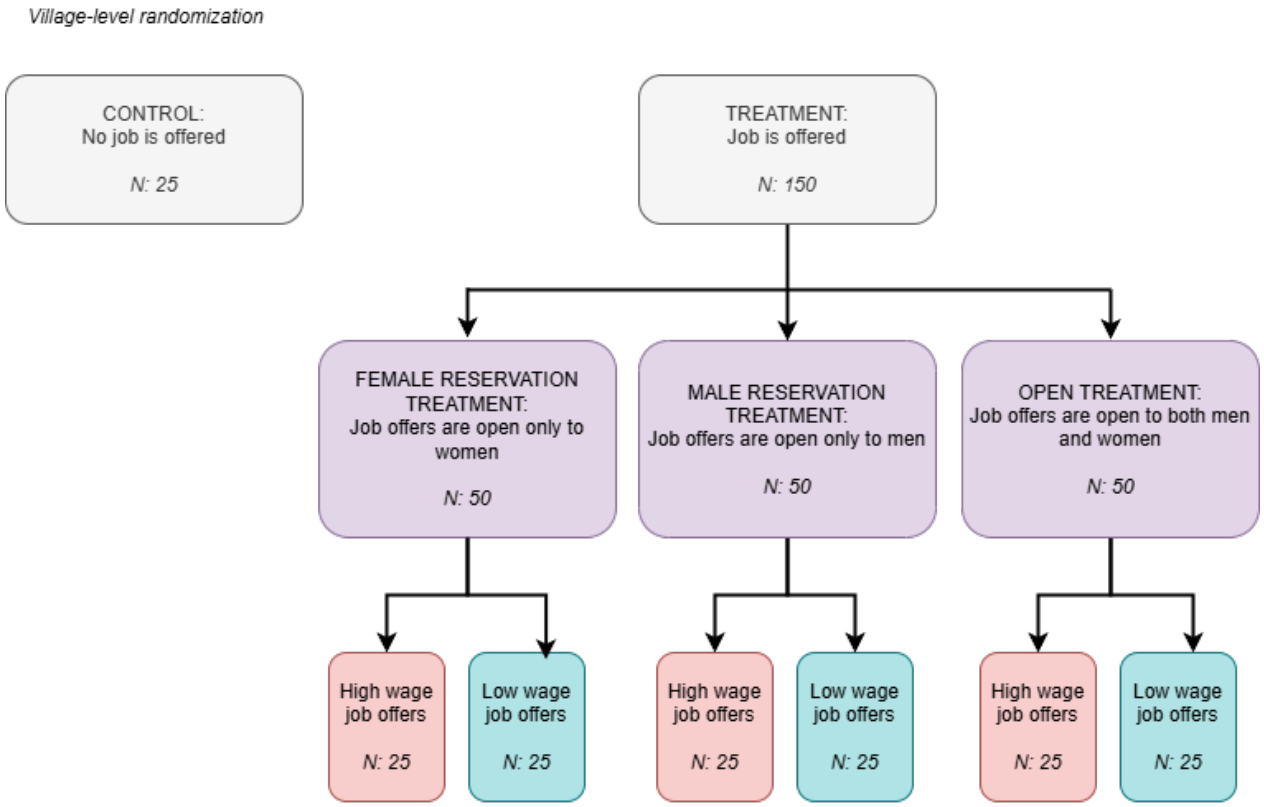
2 Experimental design

2.1 Treatment randomization

Our study is a randomized controlled trial that will operate in 175 villages of rural Bhojpur district, Bihar, India, where we offer daily wage jobs to households. Sample villages will be randomized into the following treatment arms:

1. **Control**: no job is offered
2. **Female Reservation Treatment**: job offers are open only to women
3. **Male Reservation Treatment**: job offers are open only to men
4. **Open Treatment**: job offers are open to both men & women

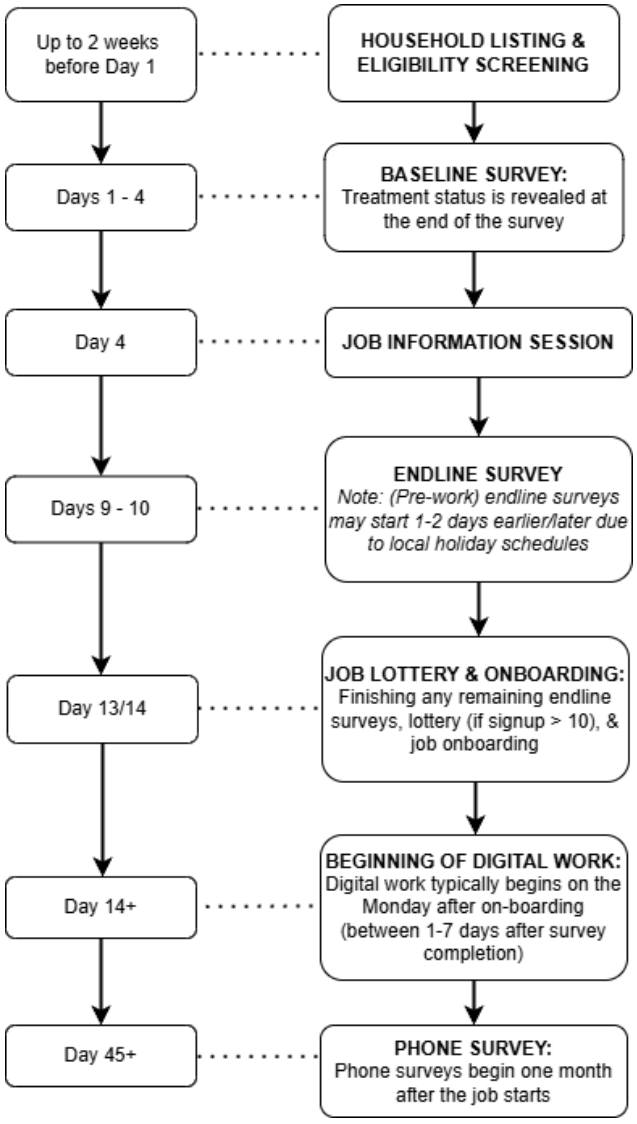
Within each treatment arm, we will cross-randomize high and low wage job offers at the village level. Communities assigned to the high wage, experimental job offers will be paid a piece rate that, if the worker completes the majority of their tasks, equals a daily wage slightly above the market unskilled wage for men; in low wage communities, job offers will earn a piece rate that, at high task completion, amounts to a daily wage below the market wage for unskilled female labor. The diagram below summarizes the randomization structure:



Note: N indicates the number of villages assigned to each arm

2.2 Study activities and timeline

The flowchart below described the timeline for each round of data collection in a village:



1. **Village selection and randomization**: Based on 2011 census data, we selected 228 gram panchayats (clusters of 3-5 villages sharing local government) where the population of the largest village had 300 or more households. Then we reached out over the phone to the gram panchayat leaders to verify information about the largest village in the gram panchayat. We removed gram panchayats where the largest village has less than 300 households, the temporary migration rate is greater than 60%, or the gram panchayat shared borders with a larger gram panchayat in the sample. From the remaining gram panchayats, we randomly selected 175 gram panchayats and randomly divided them into 5 stratas, which determined the order in which data collection will take place, i.e., one strata after another. Finally, we randomly assigned one of the 7 treatments to each village stratifying on the strata variable.

2. **Household listing and eligibility screening**: We will first visit all eligible villages to conduct a mapping survey to create a sample frame. During the listing survey, we will conduct a rapid screening exercise with the surveyed household member to see whether the household has at least one female member that meets the eligibility criteria: aged between 18 - 60 years, married, can read in Hindi, has

access to a smartphone, does not have a regular paid job, and does not have plans for temporary or permanent migration in the next one month. The household must also have at least one male member who meets the same criteria, with the exception that he may have a regular job. We conduct the listing survey with 250-300 households to ensure that we identify more than 40 eligible households.

3. **Baseline survey**: We will then visit the eligible households in a randomly determined order, with the goal of enrolling 20 households per village into the study. In the case that more than one woman is eligible in a household, we will first randomly select the female respondent. We will enroll her spouse if he is eligible, and otherwise, enroll another eligible male member, prioritizing the household head. We will double check the two respondents' ability to read simple sentences and use a smartphone and then go through the consent procedure. Once enrolled into the study, we will conduct a baseline survey covering basic demographic questions, questions on norms around women's work, and constraints faced by women related to work. At the end of the baseline survey, the respondents in the treatment groups will be presented with information about the assigned job offer as well as the treatment conditions, and invited to attend a job information session.

4. **Job information session**: Once the first part of the survey is conducted with all participants in the village, job information sessions will be conducted at a central location in the village. During the session, participants will be informed about the various features of the digital job and the app offered by our employment partner, Karya. Both enrolled households and other community members are welcome to attend this information session, where the reservation (or open) policy will be discussed again with participants. We will hold multiple information sessions within a day to provide interested community members an opportunity to attend at a time that suits their schedule.

5. **Endline survey**: Several days after the job information sessions are held, we will visit households to administer the pre-work endline survey, which will be conducted with all enrolled participants. Control households will be given the survey on the same time schedule, even though job information sessions were not held. Participants in the treatment groups will be first asked about their willingness to sign up for the Karya job. At this time, open treatment households will also clarify which eligible household member, the enrolled male or female (if any), will participate in the Karya job. Survey questions will again cover norms around women's work, phone use and perceptions around the suitability of digital work for women. The control households will hear about a hypothetical Karya job before answering the questions concerning digital work.

6. **Job lottery and onboarding**: Following the survey, all individuals from the treatment groups who have signed up for the digital work will be called for an onboarding session. We will offer work to up to 10 individuals per village. In the event that more than 10 individuals attend the onboarding session, workers will be selected through a lottery; those who are not selected for the job will leave and not be given any additional information about the job. Each individual who opts into the work and is selected for the job through the lottery (if applicable) will then be onboarded to work on the Karya app.

7. **Digital work**: 20 days of Karya work will be provided, with no work taking place on weekends and national holidays. Each worker is eligible to earn up to the maximum "wage" in his/her treatment arm, although actual payments depend on the amount of work completed by the worker.

8. **Follow-up phone survey**: After the 20 days of work are completed, we will conduct a phone survey with enrolled participants across all treatment arms to capture any additional changes in attitudes after work is completed. Along with standard survey questions, we will also use this survey to collect data on real-stakes situations for respondents to voice a preference for specific job reservation policies and to express interest in information about women's job opportunities.

3 Analysis

3.1 Hypothesis and mechanisms

Our main goal is to understand how competition for jobs (within the household) affects norms and attitudes surrounding women's work. Expressed gender norms - particularly around whether it is appropriate for women to work for pay - could change depending on whether jobs are reserved for a specific gender, or men and women are able to compete for the same job opportunities. In socially conservative communities of rural India, we hypothesize that these impacts will vary across genders, and will depend on the attractiveness of the wage, i.e., whether the offered daily wage is higher or lower than what a male laborer can typically earn in the setting.

Here we lay out several theories of norm change, based on the literature from social psychology and economics. We develop specific predictions related to each theory, against which we will benchmark our findings. This will allow us to determine which theory appears to be most in line with the observed impacts on norms, or potentially develop a new theory to explain the set of findings.

Norms do not respond to economic incentives in the short run: This is the null hypothesis that suggests norm expressions do not respond to short-run economic conditions, thereby leaving no margin to affect behaviors. In this case, treatment should be unrelated to norms outcomes.

Inference from messaging: Respondents may infer that the reservation conditions are signals that the jobs are appropriate/good for a particular gender. This means a female (male) reservation policy should make views on women's work more progressive (conservative), while the open condition should have limited effects relative to the control group. The high wage may strengthen these effects, for instance, if people interpret the high wage for female reserved jobs as signaling that women will be very good at those jobs.

Cognitive dissonance: Cognitive dissonance theory posits that when individuals are presented with two conflicting cognitions ("women should be homemakers" vs. "my wife is earning money"), this can create psychological distress, which people alleviate by modifying one of the competing cognitions or adding cognitions that reduce dissonance (e.g., "the wage is so high we cannot say no") (Harmon-Jones and Mills, 2019). We hypothesize that both the female reservation condition and the open condition will induce dissonance. Thus we expect norms under open and female reserved to be more liberal than those under male reserved and control; we likewise expect norms under female reserved to be more liberal than those under male reserved and control. We do not have a prediction for differences in norms between female reserved and open: under open, "having a choice" (as in being able to choose a man for the job) and yet choosing a woman could induce more dissonance, but few households may choose women for work. Dissonance theory also predicts that under female reserved, norms should liberalize more under the low wage condition, provided takeup under the low wage condition is equal to or greater than that under the high wage condition.

Strategic norms enforcement: Men may prefer to take the attractive jobs themselves rather than letting women from their household take up those jobs. To justify this behavior, they may adopt/enforce more conservative norms, i.e. trigger norms-based backlash. We call this behavior strategic norms enforcement. This leads to two predictions related to backlash and norms.

Backlash against gender competition: The open condition leads men to compete with women for the digital job. Hence men may agree with more conservative statements under the open treatment, especially in the high wage arm as the job is more attractive. The female reservation means that men are not directly competing with women for this particular job. Furthermore, in this case, it is more economically beneficial for the household if women take up the job. Hence if men are strongly motivated by short-term economic incentives, we will observe that norms become conservative just under the open condition.

Backlash against women's work: Strategic norms enforcement could also lead to broader backlash against women's work, including under the female reservation policy. Female reservation policy means that men are unable to even enter into the competition for the job and this may generate backlash against women's work.

For example, although the job is primarily a one-time opportunity, men may worry that supporting women's work under such conditions could lead to future cases where men cannot take up jobs. This may compel men to adopt more conservative views against women's work, regardless of whether women are allowed to work for this particular job.

Hence, strategic norms enforcement can lead to backlash against women's work both under the open policy and the female reservation policy. Backlash against gender competition is one specific case in which the short-term motivations dominate the overall effects.

Distaste for reservation: Note that backlash against female reservation policies can also happen due to other reasons, such as emotional reactions. Respondents may feel cheated/unhappy that the job is reserved for a different gender. This means the female reservation policy should make men's views on women's work more conservative, while the male reservation policy may make female views more liberal. Such emotional reactions are likely to be weak under the open condition.

3.2 Summary

The table below summarizes our preliminary hypotheses for how norms will evolve under these different theories of norms change; here "+" indicates greater openness to female work/phone use, while "-" indicates less openness. Note that men and women have different underlying economic interests and therefore are expected to have different reactions under theories related to strategic norms enforcement. Therefore our predictions for strategic norms enforcement focus on male respondents alone. This also means that we will typically examine treatment effects separately by gender in order to distinguish the theories of change.

In table 1, we lay out predictions that tie observed treatment effects (in an intent-to-treat framework) to different possible responses to competition. Comparisons here contrast treatment arms to the control group, which is not offered a job. These comparisons assume that offering a job does not have any large "level" effects (e.g. through an increase in potential income). To the extent that there is a large "new job offer" effect common to all treatment arms, the differences outlined below may not be detectable.

Table 1: Predictions for the Direction of Effects on Norms under Different Theories of Norms Change: Relative to the Control Group

			Strategic norms enforcement (male respondents)		
	Messaging Inference	Cognitive Dissonance	Backlash against gender competition	Backlash against women's work	Emotions against reservation
Female Reserved	+	+	0	0	- (male)
Male Reserved	-	0	0	0	+ (female)
Open	0	+	-	0	0

Notes: Differences are relative to the control group. + and − indicate more liberal and conservative views, respectively.

In Table 2, we instead compare treatment arms directly to the "open" treatment arm, ensuring the potential income effect is held constant across (non-control) arms. We also describe predictions for changes across wage levels within a treatment arm.

Taken together, these predictions convey our current understanding of how individuals will react to the treatments under the stated assumptions and potential responses to competition. It is possible that we will identify another underlying mechanism driving our results, in which case we would supplement the below tests with those relevant to elucidating this additional interpretation.

Table 2: Predictions for the Direction of Differences Effects on Norms under Different Theories of Norms Change: Relative to Open and Low Wage

			Strategic norms enforcement (male respondents)		
	Messaging Inference	Cognitive Dissonance	Backlash against gender competition	Backlash against women's work	Emotions against reservation
Differences relative to open					
Female Reserved	+	np	+	np	- (male)
Male Reserved	-	-	+	+	- (female)
Differences Relative to Low Wage					
High Wage: Open Condition Only	0	-*	-	np	0
High Wage: Female Reserved Condition Only	+	-*	0	np	-

Notes: * Prediction only holds under equal or greater takeup of female work in the low wage condition. np = no prediction.

Because we are starting off from the hypothesis that gender norms and expressions are malleable even in the short-run, we note the possibility that not all outcomes may move in the same direction or in ways that we foresee. For example, male respondents facing female reservation may update their views on women's phone use more favorably - as to encourage female job take-up of digital jobs - and yet strengthen their conservative views towards non-digital jobs. We will examine the outcomes as planned and interpret the results to the extent possible, thinking about the above (and other) mechanisms.

If feasible in terms of sample size, we will also leverage the pre-work lottery to assess the extent to which digital work itself shifts norms. To do this, we will examine relevant post-work phone survey outcomes for respondents who attended the job lottery, assessing results to compare those selected through the lottery to participate in the digital work to those not selected, controlling for treatment status and standard controls.

3.3 Regression specification

We will use two "baseline" regression specifications to test our hypotheses. First, the following "long" model will let us study the effect of the different reservation policies and how they interact with the cross-cut wage treatment:

$$y_{ihk} = \beta_0 + \beta_1 reservF_k + \beta_2 reservM_k + \beta_3 open_k + \beta_4 open \times high_k + \beta_5 reservF \times high_k + \beta_6 reservM \times high_k + \beta_7 female_{ihk} + \gamma X_{ihk} + \lambda_k + \epsilon_{ihk}$$

where y_{ihk} is the outcome of interest for individual i in household h in village k, $reservF_k$, and $reservM_k$ identify female and male reservation villages, $open_k$ identifies open treatment villages, $high_k$ identifies high wage villages, $female_ihk$ is a dummy variable identifying female respondents (only relevant when pooling genders), X_ihk are double lasso-selected controls, and λ_k are survey strata fixed effects. The omitted group is the control category, and (conditional on treatment) low wage. As warranted by theory-driven tests, we will also test the equality of different regression coefficients (e.g. $\beta_1 - \beta_3 = 0$. Depending on which tests are most useful for distinguishing between theories (Table 1 vs Table 2) we may also consider using the open group, low wage as an omitted group.

When testing a hypothesis does not involve interactions between the wage and the reservation policy, we

will use the following specification:

$$y_{ihk} = \beta_0 + \beta_1 \text{femR}_k + \beta_2 \text{maleR}_k + \beta_3 \text{open}_k + \beta_4 \text{high}_k + \beta_5 \text{female}_{ihk} + \delta X_{ihk} + \lambda_k + \epsilon_{ihk}$$

As before, we may also consider a version where the omitted group is the open category. In all cases, we will cluster standard errors at the village level. Finally, we note that some hypotheses and/or outcomes will not be relevant for the control group; in these cases, we will drop the control group from our regression specifications and use the open group as the omitted category. We will also test robustness of the results using an ANCOVA-style regression with baseline norms index controls, instead of the double lasso-selected controls.

4 Definitions of outcomes and control variables

One of our principal aims is understanding how and whether norms governing women's work and economic engagement are influenced by gender reservations for jobs, the prevailing wage rate, and their interaction. Our post-treatment surveys therefore devote a significant amount of time to measuring norms. Here we describe how we classify and measure norms-related outcomes.

4.1 Primary outcomes: Norms outcomes from the (pre-work) endline survey

In the (pre-work) endline survey, we ask participants about their views on various norm-related statements. The above table summarizes how we perceive each statement to relate to different themes. We try to elicit people's opinions/attitudes related to three behaviors: women's work, women's phone use, and digital work. These questions are divided into sections based on whether they are related to norms themselves (what people believe is appropriate), consequences of norm violations (costs associated with violating the norm), or the suitability of work by gender. Hence the questions are organized into seven sections.

For some questions, we ask whether respondents agree or disagree with a series of different norm-related statements. For each of such questions, we create an indicator for whether they agree with the more conservative statement. For some other statements, we elicit second-order beliefs regarding how many men in their community (out of 10) would agree with the focal statement. For such questions, we create a continuous variable that increases with conservativeness of the answer (e.g., share of men who agree with the conservative statement). To construct outcome variables, we will combine all the variables into GLS-weighted indices following Anderson (2008). Our "primary norms index" will aggregate all questions on first order beliefs (NW, CW, NP, CP, ND, and SD). We will construct a separate index to study effects on second order beliefs using questions in SO. In addition to evaluating treatment effects on this overall norms index, we will also study impacts on sub-indices that correspond to each of NW, CW, NP, CP, ND, and SD, to be examined in separate regression.

Prior to designing this research project, we conducted formative qualitative research to better understand how norms governing women's economic behavior are operationalized. While most individuals reported little "in principle" opposition to women's work, they often provided justifications or reasons that women could not work, which mapped to deeper norms tied to masculinity and femininity. More specifically, part of being a "good husband" entails being a breadwinner and economically providing for the family; having a wife who works (especially if she outlearns her husband) violates this breadwinner norm. Part of being a "good wife" is to manage domestic matters, namely caring for the home, family, and children. Individuals felt that spending too much time at work could limit women's ability to fulfill these responsibilities, or send a signal to others that they were not prioritizing these responsibilities and upholding the domesticity norm. Finally, individuals voiced concerns that travel and interpersonal interactions with others required for many types of jobs - and for navigating the internet - may risk women's safety and purity (or invite speculation about it), which is central to a family's honor. Thus, purity norms also play a major role in restricting women's economic behavior.

We therefore designed our norms statements so that some of them will relate to each of these sub-themes as

follows: domesticity norm, purity norm, and breadwinner norm. In order to better understand mechanisms, we will aggregate the questions (without the second order belief questions) into three sub-indices focused on each of these themes and report impacts on each sub-index. In an appendix, we will report results for individual questions. The list of questions and categorization are listed in the section 7 below.

4.2 Secondary outcomes

4.2.1 Post-work outcomes

The (post-work) follow-up phone survey will be conducted over the phone with all study participants to capture additional changes in attitudes after actual work is completed. It is important to note that the impacts on these outcomes may be both due to being offered specific reservation policies and the effect of participating (or having an opposite gender household member) participate in digital work and completing work offered to the household. Assuming we have sufficient interest in work for this comparison to be appropriately powered, we will leverage outcomes of the job lottery to help identify the "participation effect." Note here that the participation effect will only be identified for households who are interested in work and submit a name to the lottery.

In the phone survey, we will measure attitudes towards women's work using two real-stakes questions.

4.2.2 Preference over job reservation policy for future work

As part of the post-work survey, respondents will be informed of an opportunity to vote to influence the future gender reservation policy in their community. We will implement future work and the chosen reservation policy in a randomly chosen community. Respondents are allowed to cast a vote for their preferred reservation policy-male reservation, female reservation, or open policy. We make a random draw from the set of casted votes, so voting increases the likelihood that their preferred policy is implemented in an actual community. The first vote is free, and for the second vote, they can choose between either voting or receiving a small compensation instead of voting. Hence, we will consider three measures of preference: the outcome of the first (free) vote; willingness to cast a second vote (at a cost); and the total number of votes for a particular reservation policy.

4.2.3 Interest in information about women's work

The second measure involves taking a costly action (e.g., filling out a form) to access website-based information related to women's work opportunities. We will create different websites for each treatment condition. We will look at two outcomes: whether the household took an action to access the website, and the traffic at the website.

4.2.4 Post-work norms

Finally, we administer a small set of questions on norms around women's work, drawing on questions used in the baseline. We select the questions that are expected to be especially relevant for measuring any impacts due to being offered and engaging in digital work.

4.3 Exploratory outcomes

While the below outcomes will not help us directly differentiate between different theories of norms change, we are interested in studying them to help characterize the effects of making jobs available to citizens under different reservation and wage policies.

4.3.1 Intra-household decision-making and psychological well-being

In the (post-work) follow-up survey, we also ask questions related to women's empowerment and bargaining power, e.g., via a subset of questions highlighted in Jayachandran et al. (2022). We also ask about recent household spending, reactions towards the digital job, and psychological well-being (e.g., via PHQ-4).

4.3.2 Job interest, take-up and productivity

The job take-up outcomes do not provide clean comparisons across the treatment groups. One important reason is that the open policy allows one of two members of the household to sign up for the job, while the other conditions are restricted to one person per household. Men and women are also likely to face different constraints to engaging in digital work. Nevertheless, we plan on analyzing differences in job takeup and interest across treatment arms to characterize household decisions.

For example, we are especially interested in understanding whether female takeup (and perceived suitability) of jobs is higher when wages are lower. Under the open policy, women may be likely to take up the job when the wage is low (e.g., because the job may be perceived as being more suitable for women), while men may be more likely to take up the job when the wage is high. To understand these patterns, we record attendance at the job information session. In addition, we ask everyone to hypothetically predict for their household who would take up the job under the different wage conditions for the open policy.

We also ask qualitative questions about why individuals are (not) personally interested in the job, why a specific person was chosen for the job, and who was involved in decision-making. We also examine, conditional on sign-up, whether worker's' daily work engagement and productivity differ across treatment conditions.

In summary, we examine how lottery attendance, job take-up, hypothetical interest in the job, and productivity vary across the reservation and wage conditions.

4.4 Experimental balance

To establish experimental integrity, we will compare each of the treatment groups to the control group on key baseline variables, using our primary approach to assess treatment effects. We will also conduct a joint test of equality across all treatment arms. We will compare balance across community-level, household-level, and person-level characteristics. For person-level characteristics, we look at female and male characteristics separately.

1. Community-level characteristics
    - Total number of households
    - Distance to nearest factories
    - Distance to the district center
    - Male wage (average of agricultural and non-agricultural casual wage)
    - Female wage (average of agricultural and non-agricultural causal wage)
2. Household-level characteristics
    - Religion (Hindu indicator)
    - Caste group (SC/ST; OBC; other)
    - Household size
    - Number of adults earning income in the past 30 days
    - Household income in the past 30 days
    - Indicator: someone in household holds a community leadership position
    - Agricultural land holdings
    - Ownership of assets
    - Expenditure on food in the past 30 days
    - Expenditure on non-food in the past 30 days
    - Whether the male respondent is the husband

3. Female characteristics
    - Age
    - Education level
    - Anxiety/depression (measured by total PHQ4 score)
    - Did any paid work in the past 30 days
    - Number of days spent on any paid work in the last 30 days
    - Living with parents-in-law
    - Years of marriage
    - Number of children
    - Total income from employment in the last 30 days
    - Bargaining power index (GLS weighted)
    - Baseline norms index (GLS weighted)
4. Male characteristics
    - Age
    - Education level
    - Did any paid work in the past 30 days
    - Number of days spent on any paid work in the last 30 days
    - Baseline norm index
    - Total income from employment in the last 30 days
    - Baseline norm index (GLS weighted)

We will not report on any characteristics where 90% or more of the sample reports the same answer. For some of the outcomes involving large values (e.g., incomes, expenditures), we will top-code the values at 99%.

4.5 Control variables

4.5.1 Baseline norm measures

We decided not to use the same norms measures in the baseline and endline, given that the surveys occur close together in time, which risked anchoring effects. In the baseline survey, we instead ask participants whether they agree or disagree with 8 different norm-related statements. They are phrased differently from the endline norm questions, and six of them concern work and two concern phone use. For each, we create an indicator for whether they agree with the more conservative statement.

4.5.2 LASSO controls

To determine which variables to include in the main regressions, we will follow Belloni et al. (2014) and select control variables using post double selection LASSO. The set of potential control variables we will use include the following:

- Baseline norms: 8 indicators
- All the variables listed in the Balance Test section
    - Some of these outcomes must be constructed from more detailed questions, e.g. asset ownership (car/tractor/truck/AC, other individual assets), paid work type (causal labor, self-employment, salaried, etc.). We will include all "component" variables in our LASSO control set.
- Survey month indicators

We will recode missing outcomes to the mean (by gender, when relevant) and include indicators for missing values in the set of potential controls.

5 Heterogeneity Analysis

Our main analysis examines how expressed norms vary with reservation treatments and wage conditions. These impacts could vary with how economically attractive the jobs are perceived to be by the respondents. For example, jobs may be more attractive to men who have worse outside options. Hence, we plan to examine heterogeneity along the following dimensions:

- *Male work status*: men who have stable jobs, high income, or are busier may be less attracted to the job. We ask detailed questions about income and work status in the last 30 days and who has more time in general.
- *Family structure*: parents-in-law may also express disapproval of women taking up the digital job over her husband. Hence the effect could differ by the presence of parents-in-law in the household, or whether the male respondent is the husband.

The effects may also differ depending on how constrained women are in terms of bargaining power and access to work at baseline. We expect this section of the analysis to be more exploratory, as the predictions are not precise. We plan to examine heterogeneity in treatment effects with respect to the following:

- *Female bargaining power*: we ask questions based on a five-question survey module for women's agency in Jayachandran, Biradavolu, and Cooper (2023). We also ask about women's involvement in spending decisions and agency in taking up work.
- *Baseline attitudes*: The effects may only present for the households that have more (or less) conservative attitudes at baseline and therefore more room to update.
- *Constraints on women's work*: to understand which women are most affected by short-run norms changes, we plan to study heterogeneity with respect to reasons why women may not be working at baseline, such as family opposition, travel safety concerns, or household chore burdens. We will also consider baseline attachment to the labor market as a broader indicator of constraint status.

6 In-App RCT: Can Empowerment-Focused Stories Change Worker Attitudes?

We will also run an "in app" experiment within the subset of individuals who take up Karya work. For all workers on the Karya platform, their main task is to read sentences aloud. As an add-on intervention, we will include some stories in the sentences that Karya workers have to read as part of their daily digital work. Treatment stories will be related to female empowerment, e.g., a female protagonist who is able to work and contribute to her household. Placebo stories will be the same stories with a male protagonist, and we will include a "pure" control arm where Karya workers are assigned random sentences to read. We will randomize assignment of story workstreams at the individual level to assess whether reading the female empowerment stories over the 20 days of work directly affects norms.

For the secondary gender messaging experiment, we randomize workers into three different story conditions:

- Female empowerment stories: featuring female protagonists overcoming a problem
- Placebo stories: same stories with male protagonists
- Pure control: random sentences

We will examine impacts on worker productivity and reported attitudes. In addition to the follow-up survey data collected as part of the main RCT, we will administer a small set of questions directly through the Karya app. These questions relate to worker comprehension of stories, questions to measure gender attitudes through a conjunction fallacy approach, and job satisfaction.

$$y_{ihk} = \delta_0 + \delta_1 story_{ihk} + \delta_2 placebo_{ihk} + \delta_3 treat_k + \delta_4 female_{ihk} + \gamma X_{ihk} + \lambda_k + \epsilon_{ihk}$$

The main specification is as follows: where y_ihk is the outcome of interest for individual i in household h in village k, $story_ihk$ and $placebo_ihk$ indicate female empowerment story and placebo treatments, $treat_k$ control for the main intervention assignments, $female_ihk$ is a dummy variable identifying female respondents (only relevant when pooling genders), X_ihk are double lasso-selected controls, and λ_k are survey strata fixed effects. The omitted group is the pure control category. The sample only includes the workers who complete the sign-up process for the job.

Since we expect treatment effects to differ by gender, we will also run these regressions separately by gender (omitting the female dummy) and use a "long" regression specification that interacts the female dummy with the experimental arms to formally test for differences between genders.

7 Appendix: List of norms questions

7.1 List of baseline norms questions

Code	Relevant norms	Questions	Progressive
Baseline norms (B)			
B1	Work	When jobs are scarce in the community, men should have more rights to a job than women.	
B2	Work	When a woman earns money, she should be free to spend it however she likes.	P
B3	Work	If the household earns enough, it is better that a woman does not work for pay.	
B4*	Work	Women should never work for pay.	
B5*	Work	Women should not work for pay outside the house.	
B6*	Work	Women should not work for pay outside the village.	
B7	Phone	A family should supervise how a woman is using smartphones.	
B8	Phone	It is appropriate for a married woman to use smartphones whenever she wants.	P

- We create an indicator for agreeing with each statement, which suggests more conservative gender attitudes.
- The statements marked with P are associated with progressive attitudes, so the indicators are coded in the opposite way.
- B4-B6 are coded such that the respondents can agree with at most one statement.

7.2 List of pre-work endline norms questions

Code	Relevant norms	Second order beliefs elicited	Questions	Progressive
Norms related to work (NW)				
NW1			A woman does not need to take her family's permission to work for pay.	P
NW2			Women should work for pay only under dire financial circumstances.	
NW3			When it is difficult to find jobs, men should have priority access to work.	
NW4	Domes.		A woman's most important responsibility is taking care of the home and family.	
NW5	Purity		A woman who only does household work is of purer character than one who goes out to work.	
NW6	Bread.	SO	If a woman works for pay, it indicates that the man is not earning enough money for the household.	
Consequences of violating work-related norms (CW)				
CW1	Bread.		If a wife earns more money than her husband, it can cause domestic problems between them.	
CW2	Bread.		If a wife earns more money than her husband, it can cause domestic problems between them.	
CW3	Purity	SO	A woman working for pay risks bringing shame and dishonor to her family.	
CW4	Purity		When a woman goes out to work, other people will talk behind her back.	
CW5	Domes.		It is generally easy for women to both work for pay and fulfill their duties as a wife and a mother.	P
CW6	Domes.	SO	When a mother works for pay, children will not receive sufficient care.	
Norms related to phone use (NP)				
NP1			It is important that the man of the house monitors a woman's smartphone use.	
NP2			Mobile internet use is just as essential and useful for women as is for men.	P
NP3			Men are better equipped to deal with the risks of using mobile phones, such as scams or data theft.	

Code	Relevant norms	Second order beliefs elicited	Questions	Progressive
Consequences of violating phone-related norms (CP)				
CP1	Bread.	SO	Using mobile phones might make a woman difficult to control for her husband.	
CP2	Purity	SO	Women's mobile phone use can lead to bad interaction with other men.	
CP3	Domes.	SO	If women use the phone, they cannot concentrate on their household responsibilites.	P
Norms related to digital job (ND)				
ND1	Purity		Do you think a woman with decent character is likely to do this job?	P
ND2	Bread.		Do you think a job that pays this amount should be done by men?	
ND3	Domes.		Do you think a woman doing this job would find it difficult to maintain the upkeep of the household, in terms of cleaning, cooking, childcare, and other household work?	
Suitability of digital job (SD)				
SD1			Is this job more suitable for a man or a woman?	
SD2			In general, who will be better at doing these tasks more quickly, men or women?	
SD3			In general, who will be better at doing these tasks with fewer errors, men or women?	
SD4			In general, who will have more available time to work on these jobs, men or women?	
SD5			In general, who will be more consistent at doing the job every day for a month, men or women?	
SD6			In general, who will be more interested in this kind of job, men or women?	

- We create an indicator for agreeing with each statement, which suggests more conservative gender attitudes.
- The statements marked with P are associated with progressive attitudes, so the indicators are coded in the opposite way.
- For those in the SD section, we create an indicator for choosing men as being more suited for the job aspect.
- In addition to analyzing questions in each section as a set, we will analyze the statements marked with specific themes also as separate subsets (domesticity, purity, and breadwinner).
- For each of the six statements marked with SO, we also elicit second-order beliefs by asking the following: how many men out of 10 in your village will think that ... (statement). We analyze the second-order questions as a separate subset.