Bringing Work Home:

Flexible Work Arrangements as Gateway Jobs for Women in West Bengal *

Lisa Ho Suhani Jalota Anahita Karandikar

October 23, 2024

Abstract

There is a large latent labor force in developing countries. Several hundred million women say they would prefer to have a job and yet are out of the labor force, often because available opportunities are incompatible with traditional gender roles. In a field experiment with 1,670 households, we partner with a digital jobs platform to offer short-term data entry work to out-of-labor-force women. We find three main results. First, flexible work-from-home jobs are highly effective at bringing out-of-labor-force women into paid work. Job flexibility more than triples take up from 15% for an office job to 48% for a job that women can do from home while multitasking with childcare. Second, these flexible work-from-home jobs can also act as a stepping stone to less flexible work. The opportunity to try paid work from home increases women's take up of office-based jobs two to three months later. This opportunity to have a "gateway job" may be especially important for women from traditional households: job flexibility is more important to the labor supply of more traditional women, and work experience in turn shifts the gender attitudes of these women to become less traditional. Third, work-from-home comes with tradeoffs. It selects for less productive workers and causes workers to complete tasks more slowly and less accurately. However, these drawbacks are outweighed by the large positive extensive margin labor supply response to work-from-home. Thus, flexible work arrangements can both attract women to the labor force and provide a gateway to outside-the-home jobs.

*Ho: lisa.y.ho@yale.edu (corresponding). Jalota: suhani@stanford.edu. Karandikar: ak234@student.ubc.ca. We are extremely grateful to Esther Duflo, Ben Olken, Frank Schilbach, and David Atkin, who have advised this project from its inception, for their invaluable counsel and encouragement throughout this project. We are also very grateful to David Autor, Abi Adams-Prassl, Abhijit Banerjee, Kevin Carney, Claudia Goldin, Anatol Klass, Madeline McKelway, Lucy Page, Rohini Pande, Hannah Ruebeck, Garima Sharma, Advik Shreekumar, and audiences at the NBER Summer Institute on Gender in the Economy as well as the MIT development and labor lunches for immensely helpful suggestions. We thank Ishita Gupta, Rakesh Pandey, and Sharvari Ravishankar for excellent research assistance. We are extremely grateful to Bill Thies and Vivek Seshadri at Microsoft Research, who were exceptionally supportive and patient with us, and without whom this project could not have happened. We thank all our study participants for their time and patience. Lisa Ho thanks The Bill and Melinda Gates Foundation, awarded through the NBER, for financial support during this project. This project was funded by The Weiss Fund for Research in Development Economics, J-PAL Gender and Economic Agency Initiative, J-PAL Jobs and Opportunities Initiative, the Wellspring Philanthropic Fund via the Agency Fund, STEG, and The George and Obie Shultz Fund. We received IRB approval from MIT (#2106000402A007) and IFMR in Chennai. The experiment was pre-registered on the AEA registry (AEARCTR-0009190).

1 Introduction

Several hundred million women would prefer to work for pay but are out of the labor force.[1] This large latent workforce has high-stakes implications for both equity and efficiency: independent income increases the agency and wellbeing of women and girls, and recent studies find widespread misallocation of men's and women's labor (Hsieh et al., 2019; Ashraf et al., 2023). This misallocation, and consequently potential gains, are particularly pronounced in South Asia, the Middle East, and North Africa. Despite advancements in women's education, these regions exhibit low female employment rates compared to other areas with similar levels of economic development.

In countries with lagging female labor force participation rates, social norms are thought to be a key barrier to women's paid work (Jayachandran, 2021). The mismatch between available jobs and jobs women could do without violating norms suggests two potential strategies: (1) change *norms* so that existing jobs incur a smaller cost, or (2) change *jobs* to be more compatible with existing norms. Previous work shows that directly changing people's attitudes towards women and work is very difficult (e.g. Dean and Jayachandran, 2019), so in this project we start with the second approach: change jobs so that women can do them while incurring lower norms-related costs. If a woman's earned income and actual labor supply affects people's attitudes about woman and work, then our approach of starting with changing jobs might in turn change gender norms.

Using a field experiment in West Bengal, we test the effects of offering an emerging type of work that is relatively compatible with existing norms of women's behavior. Given the widespread expectation that women take sole responsibility for household chores, the experiment includes piece-rate data entry jobs on an online platform that allow women to work from home, multitask work with childcare, and/or choose their work hours. Jobs fitting this description have become more common, accelerated by the covid-19 pandemic.[2]

Motivated by this growth in flexible work arrangements — combined with their potential to bring previously excluded women into the labor force — we design our experiment to speak to three main research questions. First, does offering flexible work arrangements increase female labor force participation, and if so, which dimensions of flexibility are important? Second, given that many salaried jobs require in-person attendance, can flexible jobs act as a stepping stone to less flexible jobs for women initially only allowed to work from home? Third, to assess the viability

[1] This estimate is based on a 2016 poll conducted by the International Labour Organization and Gallup in 142 countries. Agte and Bernhardt (2023) estimate that more than 100 million women in India alone are disallowed from working.

[2] Online gig work opportunities are increasing in developing countries (Datta et al., 2023), and a silver lining of the pandemic is that many salaried jobs also now grant greater flexibility (Goldin, 2021). Early research suggests that these gains in flexibility are here to stay (Aksoy et al., 2022) and are having large effects in high-income countries, e.g. in the US, where labor force participation is at an all-time high for women with children under five (Bauer and Wang, 2023). In developing countries, where commuting and childcare infrastructure are less extensive, and societal norms against working outside the home are (in some cases) stronger, the impact could be even more substantial.

of employers introducing flexible work arrangements, what are the effects of flexibility on job performance and the composition of workers drawn into the firm?

We randomly assign women from 1,670 lower-middle-income households in a mix of urban, peri-urban, and rural areas to either receive a job offer or to a control condition with no job offer. Among those receiving job offers, we introduce variation along three dimensions of job flexibility: (1) the ability to choose one's work hours each day, (2) the ability to combine working with childcare, and (3) the ability to work from home. All jobs are part-time, last for one month, and are offered in partnership with Karya, a Microsoft Research spinoff company that distributes smartphone-based data tasks to gig workers in India. The job involves tasks that contribute to Bangla or Hindi datasets that can train language models, and most are copies of tasks that previous clients requested and paid for. In order to separately estimate the effects of flexible work arrangements on job performance versus worker composition, after participants have decided to accept or reject their job offer, we randomly select half of the participants who accepted a less flexible job to be surprised with an upgrade to the most flexible job (as in Karlan and Zinman, 2009). After the jobs are completed, we estimate the effects of work experience on women's gender attitudes, as well as spillovers to their children. Two to three months later, we measure subsequent take-up of different work opportunities ("jobs round 2"), including the effects of the initial at-home work experience on take up of jobs outside the home.

The primary study participants are mostly married women with little previous work experience. To focus on the extensive margin of labor force participation, to be eligible, they could not be in the labor force or enrolled in skills training. During study recruitment, to avoid selecting the sample based on interest in finding paid work, potential participants are not told that the baseline survey could lead to a job opportunity. This leads to a sample where 69% have never worked for pay prior to the study. However, to ensure women have the skills necessary to do the job if assigned to it, they must be literate in Hindi or Bangla and have access to an Android smartphone. On average, participants are thirty years old and nearly all (93%) are married. Husbands and parents-in-law play a large role in whether or not women work: only 36% of women report having the final say in their own labor supply decisions. Two-fifths of participants live with at least one of their in-laws, and 48% have childcare responsibilities for a young child under the age of eight.

We find three main sets of results. First, flexible work arrangements more than triple women's job take up. Compared to a job which requires working from an office, the most flexible job we offer — which includes the ability to choose work hours flexibly, combine work with childcare, and work from home — dramatically increases job take up from 15% to 48% ($p < 0.001$). To contextualize this 33 pp effect size, this is a larger effect on women's job take up than those found in evaluations of previous effective interventions designed to increase women's labor supply. For example, a promotional video shown to women's family members in rural Uttar Pradesh increased

job take up 78% (7 pp) (McKelway, 2023), and correcting Saudi men's second-order beliefs about women working outside the home increased job sign ups by 36% (9 pp) (Bursztyn et al., 2020).

To shed light on the mechanisms driving the effects of flexible work arrangements on women's labor supply, we randomly vary dimensions of flexibility across job offers. This allows us to separately estimate the contributions of choosing work hours, combining work with childcare, and working from home to the overall effect of flexibility on job take up. Work from home, even without multitasking work with childcare, doubles job take up from 14% to 28% ($p = 0.004$). The ability to multitask work with childcare is also important, increasing job take up by 59% ($p < 0.001$). The ability to choose work hours, however, does not make a significant difference to take up — these out-of-labor-force women can, and choose to, set aside consistent hours for paid work. Therefore, decomposing the 33 pp effect of flexibility, approximately one half (16 pp) comes from the ability to work from home and the other half (17 pp) from the ability to multitask work with childcare.

Second, we show that flexible work arrangements are not only effective at bringing out-of-labor-force women into paid work, but they can also act as a stepping stone to less flexible jobs. To test whether flexible work can act as a "gateway job" to less flexible work, we return to all study participants two to three months after the endline survey and offer them another job. The offers in "jobs round 2" vary in flexibility along the same dimensions as the initial jobs, and also introduce variation in the type of work offered. While the original jobs all consisted of online gig work, jobs in the second round also include non-digital piece-rate work (sewing masks or making jewellery) to assess whether effects on interest in work operate through digital-specific mechanisms or apply to interest in paid work more broadly. Consistent with flexible paid work acting as a stepping stone from unpaid at-home production to less flexible paid work, women who had no previous work experience before the study are 7 pp more likely to start the job they are offered 2-3 months later if they were first given the chance to experience a more flexible job during the initial intervention ($p = 0.03$). This effect is at least as pronounced when examining solely those women who were randomly assigned an outside-the-home job in round 2; they are 8 pp more likely to take up work if they were assigned a home-based job in round 1 than their counterparts in the control group ($p = 0.04$). The transition from unpaid home production as a full-time homemaker to working outside the home is a large leap — both for a woman herself and for her family members — and our results show that the short-term, flexible jobs offered in the study act as a bridge that allows women to take multiple, more manageable steps to outside-the-home work.

One mechanism consistent with this gateway effect is a mutually reinforcing relationship between pro-work gender attitudes and actual labor supply. It is women who held more traditional attitudes initially whose labor supply is most responsive to job flexibility, and in turn these more traditional women's attitudes that change most in response to work experience. The impact of

flexibility on job take up is 50% higher in households where women's pre-intervention gender attitudes are more traditional than the median participant, even conditional on other characteristics such as education, age, religion, previous work experience, cohabitation with parents-in-law, and having a young child ($p = 0.001$). Receiving a job offer in turn shifts women's gender attitudes to become less traditional by 0.05 SDs on average ($p = 0.038$), with the effect entirely concentrated on women who held more traditional pre-intervention attitudes (0.11 SDs, $p = 0.001$).[3]

In addition to the effects on the women participants themselves, there are also spillover effects on other family members' gender attitudes and participation in home production.[4] These changes could help to explain why women who first have some experience with flexible jobs are more willing to take up inflexible jobs than those without the "gateway" treatment. When the intervention ends, our survey team asks children about their attitudes and their family members' behaviors during the last month. Children with mothers in the treatment group are 16 pp more likely to say that their mother had a job ($p < 0.001$), and reassuringly, they are no more or less likely to say that their father had a job. Treatment group children also report that other household members contribute more to home production. They are 9 pp more likely to say that their fathers helped at least a bit with cooking, cleaning, or childcare ($p = 0.045$), as compared to the 42% of children in the control group who say their fathers *never* help with these activities. The gender attitudes of older children also shift to become 0.1 SDs less traditional ($p = 0.034$).

These effects on take up show that introducing greater job flexibility — even temporarily – can dramatically increase the pool of women workers that employers can hire from. So why do firms not offer flexible work arrangements more often? Beyond fixed costs and feasibility, employers would likely want to understand how adopting flexible work arrangements would affect the selection of workers drawn into the firm as well as the performance of existing employees. We separately identify selection into and treatment effects of flexible work arrangements using randomly assigned surprise upgrades to participants' initial job offers (as in Karlan and Zinman, 2009). We measure job performance using three main outcomes: (i) retention—how likely the worker is to show up to work and stay through the end of the job, (ii) quality—the average accuracy with which workers complete tasks, and (iii) efficiency—the speed with which workers complete their assigned tasks.

Flexible work arrangements improve the reliability of inframarginal workers (those who would accept the job with or without flexibility), but at the cost of quality and efficiency. The ability to work from home decreases the fraction of workers who accept the job but never start work by 25

[3]This heterogeneity is not driven by ceiling effects, as can be seen from the histograms of the gender attitudes index before and after the intervention (see Figure A.12).

[4]In the main study, the other household member who we survey is a child aged 8-18, when available. In our pilot, we also surveyed husbands, but we found that husbands who were willing to participate in our surveys were selected to have less traditional views on average, and completion rates were low (approximately 50%).

pp ($p = 0.004$). However, the treatment effect of work-from-home on productivity is negative: workers who are willing to work from the office but are randomly assigned to work from home spend more time completing the same amount of work. On the selection margin, we find that women whose labor supply is marginal to work-from-home complete tasks 7.7% more slowly ($p = 0.092$). In this context of piece rate wages, the efficiency cost is borne by the worker. However, this negative effect on productivity could pose concerns for a firm that pays according to days worked, or in a context in which productivity is more difficult to monitor and there are greater incentives to shirk.

Examining work patterns between different treatment arms, we find that flow effects can explain the negative effect of work-from-home on productivity. Defining a work session as a period of uninterrupted work in which fewer than 10 minutes passes between consecutive tasks, work-from-home causes workers to have work sessions that are 25% shorter ($p = 0.009$). In addition, women whose labor supply is marginal to work-from-home have work sessions which are 13% shorter on average ($p = 0.042$). These fragmented work patterns result in lower efficiency because of flow effects: workers complete tasks more quickly, and more accurately, when they work for a longer stretch of time without pauses. It is important to note that this does not mean that women whose labor supply is marginal to work-fromh-home have lower "ability," but rather that they are called on to switch between home production tasks and paid work more frequently.

Our paper contributes to four literatures. An extensive body of work demonstrates that female workers value flexible work arrangements more highly than men do in high-income countries (for examples, see Filer, 1985; Goldin, 2014; Goldin and Katz, 2016; Wiswall and Zafar, 2018; Mas and Pallais, 2017). These studies focus primarily on women who are in the labor force and ask to what degree compensating differentials can explain gender wage gaps. In many developing countries, however – particularly in South Asia, the Middle East, and North Africa – the first-order issue is not the gender wage gap, but rather the gender gap in labor force participation. A natural question then is to what extent women's preferences for flexibility explain low female labor force participation in these countries. In this study, we focus on women who are *not* labor market participants and show that women's greater preference for flexible work arrangements is not important just for explaining gender wage gaps, but also for explaining low female labor force participation. In addition, although women's greater preference for flexible work arrangements is well-documented in the literature, these preferences are taken as exogenous. Our experiment shows that these preferences are malleable and endogenous to women's own labor supply: work experience and a shift in gender attitudes can lessen the magnitude of these preferences for flexibility.

Second, we add to a literature studying the effects of economic behavior on gender norms. Observational studies show that different economic conditions — such as suitability to plough versus hoe agriculture — give rise to different gendered divisions of labor, and that the resulting

economic practices have an effect on gender norms in the long run (Boserup, 1970; Alesina et al., 2013). However, few experimental interventions have effectively changed gender attitudes about women and work. There is little prior evidence that norms change when women start work (with the exception of Field et al. (2021), who find that getting access to direct deposit and training increases women's labor supply and liberalizes women's own and perceived norms). In this project, we test experimentally if changing economic conditions to make them more favorable to women's employment, by increasing the flexibility of available jobs, changes gender attitudes as well as divisions of labor in the home to become more supportive of women working. Our results are promising in that working appears to kickstart a virtuous cycle for female employment, in which women working and less traditional gender roles mutually reinforce each other.

Third, we contribute to a growing literature about the effects of flexible work arrangements on job performance (Bloom et al., 2015; Choudhury et al., 2021; Bloom et al., 2022; Choudhury et al., 2022; Atkin et al., 2023; Aksoy et al., 2023; Emanuel and Harrington, 2023). Many of these papers examine trends in and impacts of post-pandemic increases in remote and hybrid work. Our project differs from previous work in our focus on the extensive margin of labor force participation, as well as our focus on developing countries. While other studies estimate which incumbent workers choose flexible work arrangements, we aim to shed light on how flexible job arrangements would change the gender composition of the labor force in a developing country context.

Lastly, we contribute to an active literature on strategies to increase female labor force participation in developing countries (Bernhardt et al., 2018; Khanna and Pandey, 2021; Field et al., 2021; McKelway, 2021; Lowe and McKelway, 2021; Agte and Bernhardt, 2023). The review paper Fletcher et al. (2017) shows that many women who are out of the labor force say that they are interested in working, and that there is a mismatch between the types of jobs available and women's job preferences. Our study shows that one important mismatch is the desire for flexible work arrangements, particularly the ability to work from home or multitask work with childcare.

The rest of the paper proceeds as follows: Section 2 describes the study population and experimental design. Section 3 presents results on the effects of flexible work arrangements on take up of work. Section 4 presents the effects of work experience on take up of future jobs, and provides evidence that the mechanisms for this treatment effect on future labor supply are changes in household attitudes and division of home production that result from work experience. Section 5 presents the impacts of flexible work arrangements on job performance, separating effects that operate through treatment versus worker selection.

2 Experimental design

We designed our experiment with three broad goals in mind. First, we want to estimate the importance of flexible work arrangements to female labor force participation, and to isolate the relative importance of different dimensions of flexibility. Second, we aim to understand what types of women are marginal to job flexibility, and to understand what the effects of flexibility would be on the job performance of inframarginal women. Finally, we aim to measure the effects of paid work experience on women and their children, including whether paid work makes women any more likely to participate in market labor in the future.

Households are recruited for the study from eight areas in and near Kolkata, West Bengal with the support of our partner NGO, Calcutta Foundation.[5] Three of these areas are rural (Canning, Noorpur, and Taldhi), three are peri-urban (Atabagan, Baruipur, and Sodepur), and two are urban (Tiljala and New Alipore). Households are eligible for our study if there is a woman who consents to participate who fulfills the following criteria: (1) she can read and speak Bangla or Hindi, (2) she has access to an Android smartphone, and (3) she is not in the labor force.[6]

2.1 Description of the Intervention Jobs

We partner with Project Karya at Microsoft Research (later Karya Inc) to implement the smartphone-based gig jobs. The jobs involve piece-rate paid microtasks to contribute to Hindi or Bangla speech datasets, which require participants to speak into their phones and record their voices speaking in Hindi or Bangla. Karya builds these and similar speech databases to better train speech recognition algorithms. One advantage of this task is that we can verify from the voice if a woman seems to be completing the tasks, even if the participant works remotely.

Tasks. Participants receive 4000 tasks to complete over the course of one month, with 1000 new tasks refreshed each week that expire after seven days. If participants are unsatisfied with their first attempt, they can re-attempt any task as many times as they wish. The tasks are presented in a fixed order, but participants can skip any and as many tasks as they desire.

There are four types of paid microtasks, which participants can choose to do in either Hindi or Bengali. The simplest tasks involve reading aloud a sentence which appears on the screen ("simple"). These sentences are selected from a database of common phone or computer commands (e.g. "set an alarm for 7am tomorrow morning"), and they are selected from tasks that previous Karya clients requested and paid for. The second type of task involves reading a sentence back-

[5] Calcutta Foundation was founded in 1994 and has worked on projects related to education, disaster relief, health, and more recently gender.

[6] We define this last criterion as (i) has not worked for pay in the last month (or spent less than 5 hours per week on any paid work) and (ii) is not in skills training. The purpose is to identify women who are housewives in the absence of the experiment.

wards, which we introduce to require more concentration ("backwards"). The third type of task involves finding a specific sentence within a paragraph, and reading that sentence out loud. The target sentence includes a specific word prompted by the task, and requires participants to locate the correct sentence within the paragraph ("paragraph"). Lastly, there are open-ended questions for participants to answer. Our partners at Karya are interested in collecting this type of speech data, as it is particularly scarce but useful ("open-ended"). Each week, the 1000 tasks are ordered in the following way: 200 simple, 100 paragraph, 150 simple, 100 backwards, 10 open-ended, 150 simple, 40 paragraph, 50 backwards, and 200 simple.

Payment. Participants earn up to one rupee per task they complete, with payments processed weekly according to task quality. Each completed task is assessed by a separate team of validators hired by our partners, who score each task on a scale of 0-2 in terms of fluency, accuracy, and volume. This means that participants could earn up to Rs 4000 (approximately $50 USD) over the course of the intervention jobs, which is equivalent to 36% of the average household monthly income in the study.

Implementing the work arrangements. The five work arrangements vary across three dimensions: time, multitasking with childcare, and location (see table below). The most flexible job we offer allows participants to work from home, at any time they choose, and while multitasking their work with childcare. In each subsequent job, we switch off one or more of these dimensions.

Work Arrangement	Time Flexibility	Multitasking with Childcare	Work From Home
1. Most Flexible	✓	✓	✓
2. Time-Inflexible	✗	✓	✓
3. Child-Inflexible	✓	✗	✓
4. Time- & Child-Inflexible	✗	✗	✓
5. Office	(✓)	✗	✗

Time Flexibility. In the time-flexible groups, participants can choose to work for as many hours as they like and at any time of day to complete the tasks before they expire. In the time-inflexible groups, participants choose a 3-hour timeslot during the job offer stage, and they can only work during that timeslot for the rest of the month. Working for three hours each day would be more than enough to complete the weekly assigned tasks.

Multitasking Work with Childcare. In the groups which allow participants to multitask work with childcare, participants are told that they can have their children next to them while they work. In the groups without multitasking with childcare, participants are told that it is *not* acceptable to have

their children next to them while they work, and participants do not get paid for tasks they submit which have children's voices in the background of the recordings.

Working From the Office. In the location-flexible jobs, participants work from home (or another location of their choice), while in the office-based group they are required to work from one of our offices. We set up 1-2 offices per treatment area, and we do not allow participants to bring their children.[7] For logistical and safety reasons, we do not keep the offices open at all times of day, but we keep the offices open at hours that participants would likely want to come to the office. The office was open between 10am-6pm six days per week while the study was running (Monday-Saturday in most areas and Sunday-Thursday in the Muslim-majority area).

2.2 Timeline and Randomization

Study implementation was staggered across eight areas, beginning in April 2022 and ending in January 2023. See Figure 1 for a flowchart describing the experimental design.

Recruitment, Informed Consent, and Baseline Surveys. Participant recruitment takes place both over the phone, using contact details provided by the Calcutta Foundation, as well as through in-person door-to-door conversations. Potential participants are not told that the study could include a job opportunity, in order to avoid selection into the experiment based on interest in work. If women are eligible and consent to participate, they complete an extensive baseline survey that covers demographics, gender attitudes, agency, technology use, psychological wellbeing, and social contacts. When possible, we also survey children aged 8-18 about their aspirations, participation in household activities, and gender attitudes.

The baseline survey also includes an incentivized job preferences elicitation. Surveyors explain that an employer in the area is looking for female workers for a month-long, part-time job, and that the jobs involve reading Hindi or Bangla sentences to help engineers make computers understand these languages. Participants are told (truthfully) that there are already many recordings of men's voices, but not enough female voices, which is why the employer is looking for female workers. Surveyors then ask participants whether or not they would accept each of the five different work arrangements for the job. Participants are told to choose carefully because their decision will have an effect on whether they will be offered a job. The work arrangements are presented one at a time, and the order of the jobs is randomly varied between participants to avoid order effects. Surveyors explain that accepting or turning down one work arrangement will not affect their chance of getting assigned to a different work arrangement, and that job assignment is decided by lottery.

[7] In Jalota and Ho (2023), we allow women to bring their children to the office to understand what the impact of requiring office-based work would be if women can multitask with childcare. Even when multitasking with childcare is allowed at the office, requiring women to come to the office cuts take up by one half.

To check comprehension, participants are asked three (of seven possible) questions to ensure they understand what the jobs involve, the different requirements of each work arrangement, and that their choice for one work arrangement would not affect their probability of receiving a different work arrangement.

Jobs Round 1: Two-Stage Randomization (Initial Job Offers and Surprise Upgrades). After the baseline survey, we randomly assign households to the control group or one of the five job groups for their initial job offers. Randomization is stratified by three characteristics: area, smartphone ownership, and whether the participant has a child under age eight.[8] If participants said during the job preferences elicitation that they would accept the job they were assigned to if offered it, then we offer them the job. This follows the "strategy method"; in effect, we randomly select one of the questions from the job preferences elicitation and implement the participant's choice. Participants can accept or reject the job offer, or else ask us to call back in a day or two if they need more time, for example in order to discuss the job offer with family members.

After participants decide to accept or reject their initial job offers, we randomly select half of the participants who were assigned to any job other than the most flexible job to be surprised with an upgrade to the most flexible job. Women who were randomly assigned to an upgrade, but who turned down the initial less flexible job, are also offered the most flexible job. We include these randomly assigned surprise upgrades, following Karlan and Zinman (2009), in order to separately identify the characteristics of women who select into flexible work arrangements from the effects of those arrangements. To estimate selection, we compare measures of job performance (e.g. quality-adjusted output) between participants who initially accepted the most flexible job with participants who were upgraded to the most flexible job after initially accepting a less flexible job. This strategy to measure selection holds constant the job that participants are doing, while comparing participants who are "flexibility compliers" (i.e. women who will only work when the job is flexible) to those who are not. To estimate treatment effects of flexibility, we compare job performance between participants who initially accepted an inflexible job and were upgraded versus those who also accepted an inflexible job but did not get a surprise upgrade. This strategy holds constant the type of worker – an always-taker on this dimension of flexibility – and assesses how their job performance changes if given greater work flexibility. After the upgrades, the jobs are implemented for one month as described in Section 2.1.

Endline Surveys and Jobs Round 2. After the randomized jobs intervention, participants complete an endline survey (within two weeks of job completion for treated participants). The endline survey takes approximately one hour to complete and covers gender attitudes, agency, and psycho-

[8] One region (New Alipore) was not stratified further for randomization, because of the small number of participating households (66).

logical wellbeing. We also survey children when possible, covering their aspirations, participation in household activities, and gender attitudes. All women completing the endline survey also take part in an incentivized job preferences elicitation. This job preferences elicitation again uses the strategy method to make participants' choices have real stakes that will be implemented as Jobs Round 2. The job choices are identical to those in the initial jobs intervention (Jobs Round 1), except for two major differences:

1. *Non-Digital Jobs.* We introduce two non-digital jobs, bringing the total number of jobs that we ask each participant to make a choice about to seven. One of the non-digital jobs could be done from home, while multitasking with childcare, and at any hours the participant wishes. The other non-digital job required participants to work from an office. The purpose of these non-digital jobs is to assess whether any treatment effects on interest in work are restricted to digital work – for example due to digital-specific skill or confidence building, or whether there are treatment effects on interest in paid work more broadly. The non-digital jobs involve constructing newspaper bags, making earrings, and sewing masks. As in the digital jobs, payment is piece rate and dependent on quality.

2. *Costly Choices.* At endline, participants do not just state whether they would accept or reject each job; instead, they choose between each job and a gift (a pressure cooker), which introduces a cost of saying yes to the job. A pressure cooker (worth approximately Rs 750, or $9) was selected as the gift because they are very common, useful household items.

We randomly select one of the seven job versus gift questions for each participant and implement their choice on that question. This means each study participant receives either a job offer or a pressure cooker after the endline survey. We compare take-up rates during Jobs Round 2 to estimate treatment effects of work experience (from Jobs Round 1) on subsequent labor supply.

2.3 Sample Description and Intervention Fidelity

This subsection will summarise participant characteristics, balance, and attrition across the experimental groups. Table A.8 presents the mean and standard deviations of important characteristics in the control and treatment groups after randomization, and Table 1 presents the same for participants who completed the endline survey.

Summary of Participant Characteristics. The average participant is 30 years old and married. Half have completed education through to 10th standard (approximately the same as 10th grade in the United States, with students aged 15-16), and 13% have at least started an undergraduate degree. Three quarters of the participants are Hindu, with the remaining coming from Muslim households. 40% of participants belong to a scheduled caste or scheduled tribe. The average

household size is 4.6 people, and 40% live with at least one parent-in-law. Three quarters of participants have a child under eighteen living in the household, and 48% of participants have a child under eight. Access to an Android smartphone is part of the eligibility criteria for the study, and 73% of women report having their own smartphone (i.e. not sharing their phone with any other household member). Average household income is Rs. 11,791 per month (approximately $142), meaning that participating household are lower-middle-income households in which women do not need to work for subsistence. The majority (69%) of participants have never previously worked for pay before starting the study. If participants were to get a job offer, only 36% report that they would have the final say in whether or not to take the job.

Balance and Endline Completion. As seen in Table A.8, randomization produced a control and treatment group which were balanced across most major characteristics of interest, although the treatment group is more likely to be Hindu (significant at the 10% level). After Jobs Round 1, 1525 households completed the endline survey, a completion rate of 91%. As seen in Table 1, attrition is balanced between the treatment and control groups, and the groups are mostly balanced on important covariates. The control group is less likely to be Hindu ($p = 0.064$) and more likely to have parents-in-law in the household ($p = 0.098$). To control for any imbalances, we present results from specifications that include covariates selected by double post LASSO in regressions estimating the effects of treatment (Belloni et al., 2014).

3 Extensive margin labor supply response to job flexibility

3.1 Measuring Job Take Up

We measure job take-up in three ways: (i) from our baseline survey, (ii) from the job offer stage, and (iii) whether participants actually start work. Each measure of job take up illuminates different parts of the decision making process about labor supply.

First, participants' baseline survey responses act as well-powered indicator of job interest. On the baseline survey, each participant is asked – for each of the five work arrangements – whether or not she would take the job if offered it. Because participants are unable to discuss the decision with family members for this measure of job take up, we interpret this baseline survey measure as an indicator of the woman's own interest in the job. Because each participant is asked about every job, one advantage of the baseline survey measure is that it is better powered to detect differences in take up between work arrangements than at later stages when each participant is offered at most one job.

Second, a participant's response when the intervention team calls her with a job offer is the measure that most closely simulates her job acceptance behavior if a firm tries to hire her. After randomization, the jobs team calls the participant with a job offer that she can respond to imme-

diately or take a couple of days to decide on, allowing her to discuss the decision with family members if she wishes. If she decides to accept the job, she is expected to begin the job as soon as possible and no later than one week from the time of her job offer. In order to make a participant's baseline survey responses meaningful, she is only contacted at the job offer stage if she indicated on the baseline survey that she would take up that job if offered it.

Third, the job take-up measure of starting work is the most real-stakes measure of labor supply: we say that a participant started work if she ever submitted job tasks to the employer for review. Participants can only start the job if they were randomly assigned to receive the job and accepted the job offer. This "started work" job take-up measure is the one we use when estimating the effects of employment, as well as the measure we use when presenting results from just one measure of job take up.

We interpret the baseline survey measure of job take-up as interest, or what the participant would like to do if she were unconstrained. During the baseline survey, she has not had time to discuss the job decision with other family members. Since only 36% of participants have the final say in their own labor supply, and nearly all participants report that they would consult with other family members before making a decision on the job offer, it is not surprising that participants' actual behavior when offered a job often does not agree with what they said they would do on the baseline survey. At the job offer stage, participants can consult with their family members if they would like to before making a decision on whether or not to accept the job. However, there is still a large gap between the rate of participants who accept jobs and the rate of participants who actually start work. We hypothesize that this difference is a combination of time-inconsistent preferences with over-optimism in participants' beliefs that they will be able to get permission to work.[9]

3.2 Estimation

Because the job offers are randomly assigned, estimating the effects of different work arrangements on job take up is straightforward. We estimate an equation of the following form using ordinary least squares (OLS):

$$\begin{aligned} \text{Takeup}_{ij} = \alpha_0 &+ \alpha_1 \text{TimeInflexible}_j + \alpha_2 \text{ChildInflexible}_j \\ &+ \alpha_3 \text{TimeChildInflexible}_j + \alpha_4 \text{Office}_{ij} + \varepsilon_{ij} \end{aligned} \tag{1}$$

where i represents a participant and j represents a job. Takeup_{ij} represents take up of job j for participant i, and is measured at three stages as described in Section 3.1. The regressors are dummy

[9] In fact, participants seem to learn from their first job offer whether or not they will actually begin work. During Jobs Round 2, treatment group participants (who have experience in receiving job offers) have greater agreement between their responses at the survey, job offer response, and starting work stages than the control group, for whom the second round of jobs is their first job offer experience in the study.

variables for each job type j, with the most flexible job excluded. When using the baseline survey measure for Takeup, we include household fixed effects and cluster standard errors at the household level because each participant is asked about each of the five jobs. We also estimate a regression in which the time-inflexible job and time- and child-inflexible jobs are combined for estimating the effects of time inflexibility, and similarly the child-inflexible job and time- and child-inflexible jobs are combined for estimating the effects of multitasking work with childcare. We find very similar results using either method.

3.3 Results

Flexibility dramatically increases women's labor supply, from 15% for an office-based job to 48% for the maximally flexible, work-from-home job. Despite the relatively comfortable conditions of the study's office-based job, only 14.6% of participants who were offered the office-based job start work. The study's office-based job offer relatively high pay, flexible hours, low commuting costs, work that was not physically tasking or dangerous, and less exposure to non-family males compared to other jobs in the area. The piece-rate wage of Rs 1 per task means that participants could earn up to Rs 4000 throughout the month, which is equivalent to 36% of average household income, while working only around 12 hours per week. The offices are open from 10am-6pm on Mondays through Saturdays, which means participants could come to the office during the most convenient time of day (mid-afternoon, as learned during our pilot study).[10] We set up 1-2 offices in each area, so in urban areas participants can walk to the office within approximately 15 minutes, and in peri-urban and rural areas participants can take short rides on an auto-rickshaw which the employer reimbursed. In addition, because the work does not involve strenuous manual labor or tasks associated with a particular religious or caste group, the job does not result in physical injuries and is not seen as belonging to an identity group incongruent with the participant's family or gender (Oh, 2023). Lastly, the participant's coworkers and managers are entirely female, shutting down concerns about harassment or other potential negative consequences of a workplace with men.

In comparison to the 14.6% take-up rate for the office job, 47.9% of women take up the maximally flexible work-from-home job, a 33 pp (or 228%) increase from the take up rate for the office-based job. This job ("Flex") allows participants to work from home, multitask work with childcare, and choose their work hours flexibly. This is a very large effect compared to other interventions that attempt to increase women's labor supply. In comparison, even very effective previous interventions, such as correcting Saudi men's perceptions of other men's beliefs about the appropriateness of women being employed outside the home, increased sign up for a job matching service by 36% (9 pp) (Bursztyn et al., 2020), and a promotional informational video shown to

[10] In Tiljala, the office was open on Sundays and closed on Fridays, as this was more convenient for the study participants in the area, 94% of whom were Muslim.

women's family members in rural Uttar Pradesh increased take up of an employment program by 78% (7 pp) (McKelway, 2023).

3.4 Mechanisms

Dimensions of Flexibility. To understand why the most flexible job increases take up by such a large amount, we separately estimate the effects of time-flexibility, childcare-flexibility, and work-from-home on job take up by comparing take up rates between the jobs which we offer in the study. Starting with the most flexible job, each job we offer toggles off one job attribute at a time, holding fixed all other characteristics.

Figure 2 and Table 2 present the take up rates for each of the five jobs, with the following main takeaways about which dimensions of job flexibility are most important to women's labor supply:

- *Time flexibility.* The ability to choose one's hours flexibly does not make a significant difference to women starting paid work. At the baseline survey stage, the ability to choose one's hours flexibly increases take up by a small amount (4-5 pp, $p < 0.001$). At the job offer and starting work stages, however, the effect of time flexibility is smaller and no longer significant. This shows that study participants — women who are part of the approximately 70% of India's female population that is not currently in the labor force — can and choose to set aside a few fixed hours most days of the week to do paid work.

- *Childcare flexibility.* The ability to multitask work with childcare increases job take up substantially. The fraction of women who start work increases by 59% (17 pp, $p < 0.001$) when women can multitask work with childcare. The effect of childcare flexibility on job take up is also large and significant at the baseline survey (28 pp, $p < 0.001$) and job offer (14 pp, $p < 0.001$) stages.

- *Work-from-home.* The ability to work from home – even at fixed hours and without multitasking with childcare – also has a substantial effect on job take up, increasing the fraction of women who start work by 100% (14pp, $p = 0.004$). This difference is also large and significant at the baseline survey (10 pp, $p < 0.001$) and job offer (10 pp, $p = 0.020$) stages.

In sum, decomposing the large effect of job flexibility on women's take up of paid work, 6% of the effect can be attributed to time flexibility, 52% to childcare flexibility, and 42% to the ability to work from home (independent of childcare concerns).

Descriptive evidence: Traditional norms amplify the importance of job flexibility. Although job flexibility increases labor supply for all participants on average, the effect is particularly pronounced among women who have more traditional attitudes about gendered household roles. This

relationship holds even conditional on other household characteristics. Traditional attitudes predict (i) lower *levels* of take up of inflexible jobs as well as (ii) a steeper *slope* in the effect of flexibility on job take up. In comparison, other household characteristics do not significantly predict the importance of job flexibility to labor supply.

To estimate the heterogeneous effects of flexibility on job take up, we use the following regression, interacting job flexibility with participant characteristics such as whether their gender-related attitudes index is more traditional than the median:

$$\text{Takeup}_{ij} = \beta_0 + \beta_1 \text{Flexible}_j + \beta_2 \text{Characteristic}_i + \beta_3 \text{Flexible}_j \times \text{Characteristic}_i + \varepsilon_{ij} \quad (2)$$

where Flexible_j is a dummy variable equal to one if j is the flexible or time-inflexible job. The results are very similar. The omitted group consists of participants who do not satisfy Characteristic_i and who are not assigned to the most flexible job offer. The coefficient of interest is β_3, which represents the differential importance of job take up if a participant satisfies the given characteristic.

Figure A.9 plots the β_3 coefficients for six binary characteristics hypothesized to be important to women's labor supply: (i) having gender-related attitudes that are more traditional than the median, (ii) not having the final say in one's own labor supply, (iii) having a young child, (iv) completing education to at least the 10th grade, (v) having higher than median household income, and (vi) cohabitating with at least one parent-in-law. The characteristics are defined so that each one is associated with a larger importance of flexibility for job take-up. The characteristic with the largest β_3 point estimate is more traditional gender attitudes ($\beta_3 = 0.17, p = 0.001$), and this is the only characteristic which has a β_3 coefficient that is significant at the 5% confidence level. Job flexibility is also substantially more important for women who do *not* have the final say in their own labor supply (significant at the 10% level). No other characteristics predict the importance of job flexibility to female labor supply. This shows that flexible work arrangements are an effective lever to include women from traditional households with lower levels of agency to participate in paid work when they otherwise would not be allowed.

Baseline gender attitudes. Table 3 and Figure 3 compare job take up rates between more traditional and less traditional participants across work arrangements grouped by their level of flexibility. Participants are classified by whether their gender attitudes index score is more traditional or less traditional than the median participant.

Measuring gender attitudes. Gender-related attitudes are measured on the baseline survey in a module with sixteen questions. Two of the questions ask participants to make hypothetical choices between two possible daughters-in-law (one who works or one who doesn't work) and between two possible sons-in-law (one who allows his wife to work and one who does not). Two questions directly ask about the participant's beliefs about whether or not women should have jobs (one about

jobs that they would do from home, and one about jobs that they would do outside the home). The remaining questions ask participants to respond on a likert scale expressing their agreement or disagreement with statements in three domains: (1) appropriate household roles for men and women, (2) women's participation in the labor market, and (3) gender and technology use.

Comparing take up rates by jobs and attitudes. Women with less traditional baseline gender attitudes have higher job take up in every work arrangement, and especially in the less flexible arrangements. As shown in Figure 3, the difference in job take up by baseline gender attitudes is smallest (5.5 pp, $p = 0.109$) for the most flexible jobs, Flex and Time. The difference is much larger for the childcare-inflexible jobs (27.8 pp, $p < 0.001$) and office job (18 pp, $p < 0.001$). The less traditional participants were 360% more likely to start the office job than their more traditional counterparts, only 5% of whom began the office job.

As can be seen most clearly in Table 3, flexibility still makes a large (+34 pp, $p < 0.001$) difference to job take up for less traditional participants, but flexibility makes an even larger difference for more traditional participants (+51 pp). The results of the regressions presented in this table control for age, marital status, work experience, cohabitation with parents-in-law, religion, number of household members, whether the participant has a child under age eight, smartphone ownership, and area.

Robustness to having a young child. To investigate whether the correlations between gender attitudes and job take up are due to a difference in likelihood of having young children, Table 3 shows the results of the same regressions interacting flexibility and attitudes for two separate subsamples: participants with and without a child under age eight. Even within both of these subsamples, participants with less traditional gender attitudes have higher job take-up. Being less traditional is associated with +19 pp (84%) higher take up of inflexible jobs for participants without young children ($p < 0.001$), and associated with +26 pp (520%) higher take up for participants with young children ($p < 0.001$). In addition, less traditional participants without children under age eight were 24 pp (67%) more likely to start the more flexible jobs than their counterparts who were more traditional ($p = 0.002$).

Agency and other characteristics. We also test whether job flexibility is differentially important to women with lower levels of agency, who have young children, who cohabit with their in-laws, who have higher levels of education, or who live in higher-income households. We find that only agency — closely correlated with attitudes — predicts elasticity of labor supply with respect to job flexibility (significant at the 10% level). This provides descriptive evidence that the importance of job flexibility for labor supply is not fixed, but is affected by traditional social norms and women's level of control over their lives.

Agency. Women who have the final say in their own labor supply are more likely to start both

the flexible and inflexible jobs, and flexibility is less likely to determine their interest in working (see Table A.10). During the baseline survey, a minority (36%) of women report that they would have the final say in whether or not to take a job. When offered a flexible job, women who have the final say are 7 pp (16%) more likely to start work ($p = 0.051$). When offered an inflexible job, women who have the final say are 15 pp (94%) more likely to start work ($p = 0.001$). These results are consistent with other work which finds women are more supportive of themselves working than their family members are (Bernhardt et al., 2018; Field et al., 2021; Bursztyn et al., 2023). In addition, our findings suggest that some of the importance of flexibility to female labor supply is due to other family members' opinions. The gap in take up between those with and without a final say in their own labor supply is smaller for the more flexible jobs, although this differential importance of flexibility is only significant ($p < 0.001$) at the baseline survey stage.

Young Children. Participants with a young child (under age eight) are less likely to start the inflexible jobs, but are equally likely to start the flexible jobs (see Table A.11). Women with young children are 9 pp (35%) less likely to start inflexible jobs ($p = 0.021$), which do not allow them to multitask work with childcare, than their counterparts without young children. Women with young children are no less likely (1 pp, $p = 0.770$) to start the flexible jobs, however. This differential importance of flexibility is positive at all stages of take up, but only significant ($p < 0.001$) at the baseline stage.

Education. More educated participants – those who have completed at least 10th standard – are more likely to start both the inflexible and flexible jobs than their less educated counterparts, and flexibility is not differentially important by education (see Table A.12). More educated women are 12 pp ($p = 0.004$) more likely to begin inflexible jobs and 19 pp ($p < 0.001$) more likely to begin flexible jobs.

Household income. Participants in higher income households are a bit more likely to start flexible and inflexible jobs than their lower income counterparts, although the difference is only significant for the flexible jobs (see Table A.13). Women in higher income households are 7 pp more likely to start the flexible jobs ($p = 0.067$) and 3 pp more likely to start the inflexible jobs (not significant). The difference in take up rates between the flexible and inflexible jobs, for women who are higher versus lower income, is small (4 pp) and not significant.

Cohabitation with in-laws. Participants who live with their in-laws are no more or less likely to start work, whether the job is flexible or inflexible (see Table A.14).

4 Gateway jobs: flexible, home-based jobs as a stepping stone

Even if flexible work arrangements are an effective lever for increasing women's labor supply, there are many jobs which are infeasible to do remotely. Work-from-home opportunities may draw

in a substantial fraction of women who want paid work but are currently out of the labor force, but if they can only ever work from home, then job opportunities for these women will remain very limited. However, work experience — including in home-based jobs — might affect the future labor supply of women who were initially constrained to only be able to work from home. The demonstration of women earning income in the home could cause the woman herself as well as her family members to update her beliefs about the costs and benefits of women doing paid work, and might also change attitudes about what "appropriate" behavior is for women.

4.1 Effects of work experience on future labor supply

Empirical strategy. To test the effects of work experience on future labor supply, we compare job take up in Round 2 between participants with different Round 1 job assignments. Because job take up is monotonic in flexibility (empirically, accepting the office job during the baseline survey more or less implies accepting the most flexible job, see Appendix Table A.24), we say that anyone whose Round 1 job assignment was more flexible than their Round 2 assignment was given a *gateway job*.

For example, if a participant was randomly assigned during Round 2 to receive the office job, then we say that she was offered a *gateway job* if she was offered the most flexible job, the time-inflexible job, the child-inflexible job, or the time- & child-inflexible job during Round 1. She was not offered a gateway job if she received the office job or no job offer during Round 1. Similarly, if a participant is randomly assigned to the child-inflexible job during Round 2, then we say she was offered a *gateway job* if in Round 1 she was offered the most flexible job or the time-inflexible job, and we say that she was *not* offered a gateway job if in Round 1 she was offered the child-inflexible job, the time- & child-inflexible job, the office job, or no job. These variables are defined analagously for the rest of the Round 2 jobs.

To test this gateway jobs hypothesis, we regress job take up (starting work) during Round 2 on whether the woman was offered a gateway job during Round 1, controlling for the Round 2 and Round 1 job assignments. We use three different models:

$$y_i = \alpha + \beta_1[\gamma_i^1 >^{flex} \gamma_i^2] + \beta_2[\gamma_i^1 \leq^{flex} \gamma_i^2] + \gamma_i^2 + \gamma_i^1 + X_i + \varepsilon_i \tag{3}$$

$$y_i = \alpha + \beta_1[\gamma_i^1 >^{flex} \gamma_i^2] + \gamma_i^2 + \gamma_i^1 + X_i + \varepsilon_i \tag{4}$$

$$y_i = \alpha + \beta_1[\gamma_i^1 >^{flex} \gamma_i^2] \times \gamma_i^2 + \gamma_i^2 + \gamma_i^1 + X_i + \varepsilon_i \tag{5}$$

in which y_i is a dummy variable equal to 1 if participant i starts the Round 2 job that she was randomly assigned to, γ_i^r is participant i's job assignment in round $r \in \{1,2\}$, and X_i is a vector of control variables selected by double lasso. γ can take on one of five values (most flexible,

time-inflexible, child-inflexible, time- and child-inflexible, or office).

In model 3, the omitted group is the Round 1 control group (i.e. participants who received no job offer during Round 1). β_1 is the effect on Round 2 job take up of being offered a gateway job (more flexible job during Round 1 than in Round 2), and β_2 is the effect of being offered an equally or less flexible job in Round 1 than in Round 2. If job take up is monotically increasing in flexibility and being offered a job has no effect on future labor supply other than through work experience in that job, then β_2 will be equal to zero. In model 4, the control group and participants who were offered an equally or less flexible job during Round 1 are pooled together. If β_2 in model 3 is equal to zero, then model 4 allows us a more well-powered estimate of β_1 as the effect of being offered a gateway job on Round 2 job take up. In model 5, we look differentially at the effect of a gateway job on the take up of different Round 2 jobs (i.e. time-inflexible, child-inflexible, and office jobs). We also examine heterogeneity in treatment effects by previous work experience, comparing women who had previously worked for pay before the study with women for whom the study jobs are the very first time they have worked for pay.

Results. Table 7 presents the effects of work experience on future labor supply. Column 1 reports estimates from model 3 and shows that being offered a gateway job increases the take up of less flexible jobs by 5pp ($\hat{\beta}_1 = 0.05, SE = 0.03, p = 0.05$) two to three months after the initial jobs intervention. In contrast, being offered an equally flexible or less flexible job in Round 1 has no effect on job take up in Round 2 ($\hat{\beta}_2 = -0.02, SE = 0.03$). Column 2 adds controls for Round 1 job assignment and shows that Round 1 job assignment has no effect on job take up in Round 2 other than through its relative flexibility with the Round 1 job.

Column 3 presents heterogeneous treatment effects by prior work experience. If the gateway job effect operates through learning from the experience of earning income, then we would expect larger effects for participants who have less previous work experience. As expected, we find that the gateway job effect is larger for women who did *not* have any work experience before the study: assignment to a more flexible job in Round 1 increases Round 2 job take up by 7 pp $(SE = 0.03, p = 0.03)$, as compared to 2 pp for women who had previous work experience before the study. There are no effects on Round 2 job take up of being assigned to a non-gateway job sequence (R1 Less or Equally Flexible Than R2) for women with or without work experience.

Lastly, Column 4 presents the results of regression 5. The coefficients on these interactions are also all positive but more noisily estimated due to smaller sample sizes. However, the coefficient on the interaction with a round 2 office-based job is still positive and significant ($\beta_1^{R2=Office} = 0.08, p = 0.04$). Given the control group's Round 2 take up rate of 19%, this means that assignment to a gateway job increases office job take up by 42% 2-3 months later. To put this effect size in context, increasing women's take up of office-based work by 8pp in similar settings is very difficult.

In Jalota and Ho (2023), we find that increasing wages by over 400% (from Rs 5000 to Rs 24000 per month) achieves approximately the same increase in office job take up among married women in Mumbai. In other words, being offered a gateway job increases the take up of office-based work by the same magnitude as increasing wages by 400%.

4.2 Mechanisms explaining gateway job effect

Work experience shifts the attitudes of women and their children to become less supportive of traditional household roles, and children report that their fathers are more likely to take part in home production activities if their mother is assigned to the treatment group. This reallocation of home production activities and change in attitudes about appropriate behavior for women might drive the effect of work experience on future labor supply.

Empirical strategy. The gateway job effect may act through impacts of the jobs intervention on women's and their families' attitudes about appropriate behavior for each gender. We estimate the effects of the job treatments using variations on the following participant-level intent-to-treat (ITT) regression:

$$y_{is} = \gamma_s + \gamma_{ITT}\text{JobOffer}_{is} + \gamma_3\theta_{is} + \varepsilon_{is} \tag{6}$$

where y_{is} is the relevant outcome variable (e.g. gender-related attitudes or agency); γ_s is strata (s) fixed effects; JobOffer_{is} is the randomization into receiving a job offer; and θ_{is} is a vector of control variables selected using a double LASSO (least absolute shrinkage and selection operator) (Belloni et al., 2014). In most specifications, we pool together all job offers and compare participants who received any job offer to the control group. In some specifications, we estimate the effects of being offered each type of job separately to understand if there are different effects of being offered a flexible versus inflexible job. However, these regressions will be less well powered. The coefficient of interest is γ_{ITT}, which represents the effect of receiving a job offer.

We also report two-stage least squares (2SLS) estimates, for which we instrument job take up with being assigned to a treatment group. The first stage equation is

$$\text{JobStart}_{is} = \pi_s + \pi_{FS}\text{JobOffer}_{is} + \pi_3\theta_{is} + u_{is} \tag{7}$$

and the second stage equation is

$$y_{is} = \beta_s + \beta_{2SLS}\widehat{\text{JobStart}}_{is} + \beta_3\theta_{is} + e_{is} \tag{8}$$

where the coefficient of interest β_{2SLS} is the effect of work experience on the outcome of interest.

For the exclusion restriction to hold, the job offer needs to affect the outcome only through work experience.

Measuring changes in attitudes and home production. The endline survey gender-related attitudes index is composed of 15 questions, which are similar to those on the baseline survey. Respondents are asked to rate whether they strongly disagree, disagree, agree, or strongly agree with each of 15 statements across four domains (household roles, employment, technology use, and ability). In the main specifications, the gender attitudes index is constructed following the strategy outlined in Anderson (2008). The effects of work experience on attitudes are similar when we use other methods of aggregating the gender attitudes questions, such as taking a simple average (see Table A.16).

At both the baseline and endline survey, we also attempt to speak to one child aged eight to eighteen in each household. Children are asked about their parents' labor supply, their gender attitudes, how often they and their father helps with household chores, and their educational aspirations. The gender attitudes module is combined into an index, following the same procedure as for adult women. Questions about attitudes were similar to those asked to mothers, with some statements that might be more relevant to school children (e.g. "girls are equally intelligent as boys" or "it is more important for boys to go to university than girls.").

Effects on women's attitudes. Starting with the primary participant (the adult woman), as reported in Figure 5 and Table 4, the jobs intervention shifts women's gender attitudes to become less traditional. Assignment to a job offer group causes women's gender attitudes to become 0.05 SDs less traditional on average ($p = 0.038$). Assuming the job offers affect gender attitudes only through work experience, having a job shifts gender attitudes to be less traditional by 0.12 SDs (2SLS estimate).

The treatment effect on gender attitudes is concentrated entirely on participants whose views were more traditional at baseline (results reported in Table A.19 and Figure 5). Job offers shift gender attitudes of more traditional women by 0.11 SDs on average ($p = 0.001$), translating to a 2SLS job experience effect of 0.31 SDs. In contrast, the estimated treatment effect for women who were already less traditional at baseline is small (-0.02 SDs) and insignificant ($p = 0.557$). Figure A.12 plots the gender attitude distributions before and after the intervention for women whose baseline attitudes were more traditional versus less traditional, and shows that the lack of effect on women were who already less traditional is not due to ceiling effects.

The effect of work experience on *more traditional* women's attitudes is important because this means that job flexibility and work experience are important to the same subset of participants: women from more traditional households are the ones whose labor supply is most responsive to job flexibility, and in turn, it's these same women from more traditional households whose attitudes

change most in response to work experience. This means that flexible work arrangements are effective at targeting the labor supply of women from more traditional households, since they are the flexibility compliers and also the ones whose attitudes are most affected by work experience.

In order for this mechanism to be effective — flexible work arrangements acting as a gateway job to outside-the-home work — experience in the home-based jobs (which are the only jobs that women from more traditional households are initially willing to take) must be enough to change attitudes. Consistent with this mechanism, the effect of work experience on gender-related attitudes is driven largely by the more flexible jobs (most flexible and time-inflexible), as seen in Table 4. Being offered a flexible job shifts women's gender attitudes to become 0.07 SDs less traditional on average ($p = 0.007$), which translates to a 0.13 SDs effect of work exerience according to the 2SLS estimate. The treatment effects on attitudes of being assigned to any other, less flexible job arrangement is not significant, which could be due to lower job take up in the less flexible groups. We can reject that the being offered a flexible job has the same effect as being offered a childcare-inflexible job ($p = 0.048$), although we cannot reject that being offered a flexible job has the same effect as being offered an office job ($p = 0.231$). However, if paid work experience only affects the attitudes of women who held more traditional attitudes initially and those women were not willing to do the office job in Round 1, then we might not expect much effect of the office job on attitudes in any case.

Effects on children's attitudes. As shown in Table 6 (column 1), children whose mothers were assigned to the treatment group are 16 pp more likely to say that their mother had a job during the last month ($p < 0.001$). As a placebo check, we also ask children if their fathers had a job, and treatment has no effect on this outcome (column 2). Age and gender do not predict whether children take note that their mother has started a paid job. This reassures us that many children notice that their mother is doing paid work and so may update their attitudes in response.

As reported in Table 6 (columns 3-5), treatment shifts the attitudes of older children to become less traditional. The overall effect of treatment on attitudes of children of all ages has the same point estimate as for adult women (0.05 SDs), but the sample size is smaller, and the estimate is not significantly different from zero. Splitting the sample by age, however, treatment has a significant impact on the attitudes of children older than the median (age >12), who become less traditional by 0.11 SDs ($p = 0.034$). On the contrary, younger children's attitudes remain unchanged. This difference in effects by age may not be surprising; as Dhar et al. (2022) note, adolescence is hypothesized to be a particularly important time for morality and identity formation, as adolescents are mature enough to contemplate nuanced questions about gender role in society, while still being young enough that their views are relatively malleable (Kohlberg, 1976; Markus and Nurius, 1986). To put this effect in context, this is a very large effect on attitudes given the short-term nature of

the intervention and the lack of direct focus on attitudes in the intervention.

Examining differences by gender, boy children have attitudes that are 0.13 SDs less gender-equal than girl children on average ($p = 0.062$), showing that differences in gender-based preferences and beliefs start at a young age. However, the treatment does not affect the attitudes of boy children differentially from girl children, and so the gap in average gender attitudes between boys and girls is the same in the treatment and control group at endline ($p = 0.935$).

Effects on home production. Work experience might also affect women's future labor supply if it results in a reallocation of home production activities that makes future job offers more feasible. Consistent with this causal channel, older children assigned to the treatment groups report helping more with cleaning. As reported in Table A.25, while control group children report on average helping with cleaning on fewer than half of the days within the last month, the average treatment group child reports helping with cleaning on half of the days in the last month ($p = 0.064$). Younger children, however, do not report helping more with any chores when their mothers are assigned to the treatment group.

Examining differences by child gender, in the control group, girl children are more likely than boy children to report helping with cooking ($p = 0.016$) and cleaning ($p = 0.013$). Boys on average say that they help with cooking and cleaning between never and less than half the time, while girls report helping with both somewhere between less than half and half the time. However, similarly to gender attitudes, there is no differential effect on boy versus girl children, and so this gap remains in the treatment group. There is no treatment effect of women working on the gender gap in home production between girl and boy children.

In addition to asking children how often they help, we also ask children to what degree their fathers help with cooking, cleaning, and childcare. As reported in column 4 of Table A.25, treatment group children are 9 pp more likely to report that their fathers helped at least occasionally (more than never) during the treatment period ($p = 0.045$). Given that only 58% of children in the control group report that their fathers *ever* help with cooking, cleaning, or childcare, the job offers increase men's propensity to contribute to home production by 16%.

Discussion of effects on children. When their mothers begin working, children notice. Older children are more affected by their mothers receiving a job offer: they are more likely (i) to report helping with cleaning, (ii) to say that their fathers are helping with cooking, cleaning, and childcare, and (iii) to hold less traditional gender attitudes.

The effects on children could be driven either by the job offer or by their mothers doing paid work. Watching their mothers earn income for the household and require non-family time might change children's perceptions of their mothers. Independently of seeing their mothers work, children could infer from a job offer – if their mother discusses the job offer at home – that their

mothers' time is more valuable than they previously believed, or that their mothers are more intelligent or able than they previously believed. This could affect their willingness to help with home production, as well as their attitudes about the relative abilities of men and women. Similarly, women's husbands might learn from the job offer that their wives have better outside options than they previously believed, which could increase women's bargaining power in the household and cause their husbands to do more housework.

Ruling out increased agency as a mechanism for gateway job effect. Treatment had no significant effect on the women's agency index, which may not be surprising as the job was described as a short-term, one-time opportunity. This means we should not expect the earned income to change bargaining power within the household. To measure agency, we construct an index from six questions: (a) who has the final say in whether the participant could take a job, (b) whether the participant asks for permission before purchasing clothes, (c) whether the participant asks for permission before going out of the house, (d) whether the participant asks permission before meeting friends, (e) whether the participant's opinion is taken into account in significant purchases, and (f) whether the participant gets the final say in significant household purchases. As reported by Table A.20, there was no effect of treatment on agency when all work arrangements are pooled together, and in Panels B and C, we test for heterogeneous treatment effects on agency by work arrangement or by baseline agency and do not find any effects.

The lack of change in agency combined with women's increased willingness to work outside the home provides indirect evidence that *men* changed their beliefs and attitudes. Men are still most likely to have the final say in their wife's labor supply at the end of the study (in more than 60% of households), which is not significantly different from the fraction we measure at baseline. Given that men's preferences are still binding on women's labor supply in most households, and that treated women are more likely to work outside the home during Round 2 than in Round 1, the null effect on agency suggests that treatment did affect men's beliefs or attitudes to become more supportive of women working outside the home.

4.3 Effects of work experience on wellbeing.

In addition to any effects on future labor supply, in order to assess the welfare effects of increasing the availability of flexible work arrangements, a policymaker would want to understand how these policies would affect women's wellbeing. If women are now working for pay and continuing with most of their home production responsibilities, then the work could be come a "second shift" that reduces women's welfare, even though the women themselves say that they want to work for pay. The increase in job take up between Round 1 and Round 2 is a revealed preference measure that women enjoy their jobs, but we also measure psychological wellbeing to estimate effects on

women's welfare.

We find that work experience has no effects on the psychological wellbeing index, although work experience increases women's perceptions that their potential and talents are put to good use (see Table A.21). There is no effect on the overall psychological wellbeing index as well as the four index components individually: (i) sleeping peacefully, (ii) feeling overwhelmed, (iii) feeling happy, and (iv) feeling worried. In this module, participants are asked to rate whether over the last month (i.e. during the treatment period), they felt this way never, a few days, around half the days, more than half the days, or nearly every day. At baseline, the average participant has moderately high levels of psychological wellbeing. The average participant sleeps peacefully most days, feels overwhelmed occasionally to half of the time, feels happy around half to most days, and feels anxious occasionally to half the time. Examining each work arrangement separately, we do not see important heterogeneity by the work arrangement (Panel B). There are also no heterogeneous treatment effects by participants' baseline wellbeing (Panel C).

However, treatment participants are more likely to believe that their potential and talents are put to good use. In the control group, more than one quarter of participants say that their potential and talents are put to good use "not at all." In comparison, the treatment group is more likely to believe that their potential and talents are put to good use "a little bit," "somewhat," or "very much" (see Figure 6). Even in the treatment group, however, fewer than 15% of participants say that their potential and talents are put to good use "very much." Coding "not at all" as 0, "a little bit" as 1, "somewhat" as 2, and "very much" as 3, treatment increased participants' self-reported ability/talent use score by 9.5% ($p = 0.021$). This corresponds to an 8 pp (32%) decrease in participants who said that their talents/abilities were put to use "not at all" ($p = 0.001$). Examining differences between treatment participants who earned income and those who never started work, we can confirm that this increase in self-assessed use of talents and abilities is driven by the work who take the job (see Appendix Figure A.25).

5 Effects of flexibility on worker performance

Section 3 shows that introducing flexible work arrangements would greatly increase female labor supply. So why do employers not offer home-based work options more frequently? In addition to the effect of flexible work arrangements on women starting work, employers likely want to understand the effects of flexible work arrangements on worker performance. Using the randomly assigned surprise upgrade to the most flexible job, we estimate the effects of job flexibility on worker performance, separating effects that operate through treatment effects of flexibility versus changes in worker selection. By holding the work arrangement constant and comparing women who were willing to accept an inflexible job with those who were not, we can characterize the *flexi-*

bility compliers (women who will only accept jobs when they are flexibles. By holding worker type constant — looking only at those who accepted an inflexible job — and comparing the performance of those who stayed in the inflexible job with those who were randomly selected for an upgrade to the flexible job, we estimate the effects of flexibility on job performance for *inframarginal* women workers who are willing to work inside or outside the home.

5.1 Measuring job performance

We use three main measures of job performance: (i) reliability, (ii) output, and (iii) efficiency.

Worker **reliability** is measured by whether or not the worker shows up to the job after accepting it. In most specifications, we masure reliability as a dummy variable equal to 1 if the participant accepted the job and later started work (i.e. submitted tasks for the employer to review) and equal to 0 if the participant accepted the job but never started work. We call these participants who accept the job but never start work "no shows." No shows could be costly to the employer, because recruitment and onboarding are costly in terms of time and wages paid to the firm's hiring officers. If the participant turned down the job, she does not enter into this measure of job performance.

Our main measure of output is **quality-adjusted output**, defined by the number of tasks completed by the worker, adjusted for quality metrics defined by our partners at Karya. Because workers are paid according to quality-adjusted output, this measure is also equal to earnings. If a worker submitted a task to Karya that achieved a full score in both accuracy and volume, she was paid 1 rupee for the task. If she did not complete the task satisfactorily, then pay was lower according to the following formula: if the task submisson was not accurate, then she received 0.5 rupees, and if the task was accurate but the volume of her voice recording was not sufficiently loud, then she received 0.75 rupees. The accuracy and volume scores were determined by a separate team of validators who listened to each task. We calculate the effects of flexibility on quality-adjusted output per worker in two ways (i) only including participants who started work, and (ii) including all workers who accepted the job, whether or not they were reliable (i.e. started work).

Efficiency is measured as quality-adjusted output produced by a worker per minute. We measure minutes worked using the timestamps provided to us by Karya. During any given minute in the intervention month, we counted that minute towards time worked if the participant took any action during that minute (e.g. starting a task, listening to one of their tasks to check it for accuracy, or submitting a task). For this measure of job performance, we only include participants who started work.

5.2 Selection into flexible work arrangements

Estimation. We call women who take up jobs only when they are flexible *flexibility compliers*. To estimate the characteristics of flexibility compliers based on their job performance, we compare two groups of participants: (1) those who are initially offered and accept the most flexible job, with (2) those who were initially offered and accepted a less flexible job, but are then randomly selected for an upgrade to the most flexible job. We use the following regression to characterize the differences between these two groups:

$$y_i = \beta_0 + \beta_1 \mathbb{1}\{\text{InitialOffer}_i = \text{LessFlexible}_j\} + \varepsilon_i \tag{9}$$

where y_i is a measure of job performance, and the comparison group is participants who are initially offered and accept the most flexible job. The coefficient of interest, β_1, is the difference in job performance between participants who were randomly selected for an upgrade to the most flexible job, with those who were initially randomly assigned to the most flexible job. Comparing these two groups allows us to assess whether participants who are willing to work in a less flexible job perform systematically differently from flexibility compliers.

No selection in who starts work. Flexibility compliers are no less likely to start work after accepting the job (Figure A.13). 67% of participants who accept the most flexible job begin work, which is similar and statistically indistinguishable from the rates of starting work for participants who accept a less flexible job and then are upgraded (initially accepted time-inflexible: 66%; child-inflexible: 70%; time- & child-inflexible: 66%; office: 70%). These similar rates of starting work conditional on accepting the job across different treatment groups means that we can interpret any differences between these groups in terms of their task completion as being driven by worker composition rather than the probability of starting work.

Quality-adjusted output and hours worked. Flexibility compliers are marginally less efficient (in terms of quality adjusted output per minute) than workers who would accept less flexible work arrangements. Workers who initially accepted the most flexible job produce 0.17 Rs less output per minute than workers who were upgraded from a less flexible job ($p = 0.106$).

5.3 Effects of flexible work arrangements

Estimation. To estimate the effects of work arrangements, we compare the job performance between two groups who were initially assigned to a less flexible job offer: (1) those who accept the less flexible job offer, and (2) those who accept the less flexible job offer, but are randomly selected for an upgrade to the most flexible job. We use the following regression to estimate the effects of

the upgrade to the most flexible job:

$$y_{is} = \beta_s + \beta_1 \mathbb{1}\{\text{Upgrade}_i = 1\} + \theta_{is} + \varepsilon_{is} \tag{10}$$

where y_i is a measure of job performance, and the comparison group is participants who initially accept a less flexible job and are not randomly selected for an upgrade. We estimate this regression separately for each of the less flexible jobs: Time-Inflexible, Child-Inflexible, Time-&-Child-Inflexible, and Location-Inflexible. The coefficient of interest, β_1, is the difference in performance between participants who were upgraded and participants who were not — in other words, β_1 is the impact on job performance of relaxing the relevant dimension of flexibility.

Starting work. Working from home greatly increases the fraction of workers who start the job after accepting it, while time flexibility and the ability to multitask work with childcare do not make a significant difference to starting work (see Figure A.13). Workers who are upgraded from the office to work-from-home are 25 pp more likely to start work ($p = 0.035$). On the other hand, workers who are upgraded from time-inflexible to most flexible are no more likely to start work. Participants who are upgraded from child-inflexibility and time- and child-inflexibility appear to start work more often overall (by 6 pp and 11 pp, respectively), but neither of these effects are significant, even when the child- and timechild-inflexibility are pooled together. Taken together, this means that the upgrades provide us with a clean measure of the impact of time flexibility on job performance, but that we may need to compute bounds in order to separate worker selection from the impacts of flexibility for groups involving childcare inflexibility and office-based work. In these bounding exercises, we will assume the difference in starting work is driven by the worst performing workers (following Lee, 2009).

Quality-adjusted output and hours worked. Working from home decreases worker speed, but does not affect total quality-adjusted output. Working from home decreases the average quality-adjusted output produced per minute by 19% ($p = 0.028$) while keeping total quality adjusted output produced over the month-long job the same. However, as discussed above, given the different rates of starting work for workers who were upgraded from the office to work from home, this difference in hours worked could be driven by either worker composition or treatment effects of work. In order to isolate treatment effects alone, we compute Lee bounds and find that, even assuming that the would-be no-shows in the upgraded group are the least productive workers, working from home results in the same quality adjusted output per minute. This makes it likely that work-from-home has a negative treatment effect on efficiency. Flexibility in time and childcare, on the other hand, do not have a significant effect on total output or productivity: there is no effect of time flexibility and childcare flexibility on quality adjusted output or time spent on work.

5.4 Mechanisms: workflow interruptions

The results on job performance show that working from home increases the time required to do the same quantity of work, and suggests that flexibility compliers may be less productive than their counterparts who are willing to accept less flexible jobs. What about participants' work patterns explains these differences in job performance?

One of the major differences between working at home and working from an office is the possibility for interruptions. At home, other family members may call on participants for their immediate attention, while at the office, participants are less likely to be called away from their work. To investigate this channel, we group participant's tasks into worksessions in which consecutive tasks are completed at most ten minutes apart (following Adams-Prassl and Berg, 2017). Consistent with this hypothesis, working from the office causes a large increase in the concentration of work, in the sense that worksessions are longer and more tasks are completed in each worksession. As reported in Figure 7 (Panel B), participants who work from the office complete 68% more tasks per worksession on average ($p < 0.001$).

These interruptions matter because of task-switching costs. We find that there is a "warm up" period associated with each work session during which workers complete their tasks more slowly and with more mistakes (Figure 7, Panels C and D). Conditional on worksession length, initial tasks within a worksession are more likely to be completed incorrectly (meaning that participants will not receive full pay, and the output is not useful to the employer). And not only are these tasks more likely to be marked as inaccurate, tasks completed at the start of a worksession also take longer for the worker to complete.

Interruptions to workflow may also explain why flexibility compliers have lower productivity than women who are willing to accept inflexible jobs. Flexibility compliers have shorter worksessions than participants who say that they are willing to work from an office, which may reflect that participants who are *not* willing to work from the office know that their household responsibilities will prevent them from working in a focused way for a sustained interval of time.

In this piece-rate context, the lower efficiency of home-based workers is unimportant to an employer. However, if this negative effect of work-from-home on productivity exists even in this context where the *worker* pays the price of lower efficiency, then we might expect the negative effect to be exacerbated in salaried jobs in which the *employer* bears the cost of lower efficiency. In addition, if many of the home-based workers need to be able to switch from their job to household responsibilities many times throughout the day, and potentially unpredictably, the way in which female workers are "on call" at home may be a barrier for jobs that require teamwork or quick assignment turnarounds. The fragmented work patterns caused by home-based working arrangements and that are required for flexibility compliers might discourage employers from offering

greater job flexibility.

6 Conclusion

Many women who would like to work for pay cannot do so because available jobs are incompatible with their household roles. In a field experiment with 1,670 households in West Bengal, we study the consequences of shaping work arrangements to accomodate expectations of women's domestic responsibilities. We randomly assign women to receive one of five jobs that vary along the ability to (i) flexibly choose work hours, (ii) multitask work with childcare, and (iii) work from home, and we estimate the effects of these attributes on job take up. To separately identify the effects of flexibility on worker composition and job performance, we use a surprise job offer upgrade design similar to Karlan and Zinman (2009). Jobs are implemented over the course of one month, and a post-job survey measures effects on the gender attitudes of women and their children. Two to three months after the initial randomized controlled trial, we offer another set of jobs to participants to assess whether work experience increased future interest in work.

We find three sets of results. First, flexible work arrangements are very effective at increasing labor supply for women, particularly those from traditional households. Varying different dimensions of job flexibility, we document that the the effect of flexibility on labor supply is driven by the ability to multitask (combining work with childcare) and to work from home. These are the deciding factor in whether or not to work for many women. Second, we ask why employers do not offer flexible work arrangements more often, and we find that job flexibility decreases worker performance. Working from home causes women to work more slowly and to make more mistakes, and the women drawn into the firm by flexible work arrangements are also slower workers. In both cases, it appears that differences in work performance are driven by interruptions to workflow. Third, flexible work arrangements act as a stepping stone to less flexible jobs, including outside-the-home work. Job flexibility makes the biggest difference to the labor supply for women from more traditional households, and experience with flexible jobs in turn shifts the gender attitudes of these women and children from traditional households to be less supportive of traditional household roles. Our results highlight that there is a mutually reinforcing relationship between women's actual employment and gender attitudes that support women's work.

One implication of the gateway jobs finding is that a gradual approach to transitioning women from unpaid home production to market labor, through intermediate "stepping stone" jobs, could be effective. However, in order for this approach to not trap women in lower-paying, at-home jobs, the intermediate jobs may need to be temporary so that they do not become an absorbing equilibrium (Gulesci et al., 2023). In our study, we only offer women one job at a time: their choice is always to take the job or not have any job at all. This raises a question for future research: if women

are given the option to continue working from home indefinitely, is it possible that flexible work arrangements could result in a more gender-segregated labor market? If so, what would help to ensure that women do not get "trapped" in jobs that are more flexible but also more precarious or less well paid?

One policy implication of these findings is that offering flexible work arrangements would likely be an effective strategy for the recruitment and retention of female workers. If firms have work that can be completed from home, then it could be in the firm's best interest to allow workers to work from home, as this increases the number of potential workers they could access. However, the negative effect of work-from-home on productivity might discourage firms from offering greater job flexibility. That said, as we document in Jalota and Ho (2023), women's labor supply is very inelastic, and according to our estimates it should be possible to offset the negative effect on productivity with lower wages.

Even taking this negative effect on productivity into account, however, there are at least two reasons that work-from-home arrangements may be *underprovided*. One is that employers may lack the information or skills to offer flexible work arrangements at the efficient level. Employers may not know the benefits of remote work, for example not realizing how large the effects would be on their pool of potential workers. Employers may also be interested in introducing greater job flexibility, but may not know how to introduce flexible work arrangements without sacrificing worker performance. In addition, employers may not internalize the positive effect of home-based work experience on women's future labor supply in outside-the-home work. Better understanding firms' decisions to offer flexible work arrangements in developing countries — the labor demand side — may be a fruitful area for future research.

References

Adams-Prassl, Abi and Janine Berg, "When home affects pay: An analysis of the gender pay gap among crowdworkers," *Available at SSRN 3048711*, 2017.

Agte, Patrick and Arielle Bernhardt, "The Economics of Caste Norms: Purity, Status, and Women's Work in India," 2023.

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate, "Working from Home Around the World," 2022.

_ , _ , _ , Steven J Davis, Mathias Dolls, and Pablo Zarate, "Time savings when working from home," Technical Report, National Bureau of Economic Research 2023.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the origins of gender roles: Women and the plough," *The quarterly journal of economics*, 2013, *128* (2), 469–530.

Anderson, Michael L, "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*, 2008, *103* (484), 1481–1495.

Ashraf, Nava, Oriana Bandiera, Virginia Minni, and Victor Quintas-Martinez, "Gender roles and the misallocation of labour across countries," *Working Paper*, 2023.

Atkin, David, Antoineete Schoar, and Sumit Shinde, "Working from Home, Worker Sorting and Development," *Working Paper*, 2023.

Bauer, Lauren and Sarah Yu Wang, "Prime-Age Women Are Going Above and Beyond in the Labor Market Recovery," *The Hamilton Project*, 2023.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 2014, *81* (2), 608–650.

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer-Moore, "Male social status and women's work," in "AEA Papers and Proceedings," Vol. 108 2018, pp. 363–67.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying, "Does working from home work? Evidence from a Chinese experiment," *The Quarterly journal of economics*, 2015, *130* (1), 165–218.

_ , **Ruobing Han, and James Liang**, "How hybrid working from home works out," Technical Report, National Bureau of Economic Research 2022.

Boserup, Esther, *Woman's Role in Economic Development*, George Allen and Unwin Ltd, 1970.

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

_ , **Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott**, "How Are Gender Norms Perceived?," 2023.

Choudhury, Prithwiraj, Cirrus Foroughi, and Barbara Larson, "Work-from-anywhere: The productivity effects of geographic flexibility," *Strategic Management Journal*, 2021, *42* (4), 655–683.

_ , **Tarun Khanna, Christos Makridis, and Kyle Schirmann**, "Is hybrid work the best of both worlds? Evidence from a field experiment," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, 2022, (22-063), 22–063.

Datta, Namita, Nadina Iacob Natnael Simachew Nigatu Mpumelelo Nxumalo Rong Chen (with Sunamika Singh Clara Stinshoff, and Luka Klimaviciute et al.), "Working Without Borders: The Promise and Peril of Online Gig Work," 2023.

Dean, Joshua T and Seema Jayachandran, "Changing family attitudes to promote female employment," 2019, *109*, 138–142.

Dhar, Diva, Tarun Jain, and Seema Jayachandran, "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India," *American economic review*, 2022, *112* (3), 899–927.

Emanuel, Natalia and Emma Harrington, "Working Remotely or Remotely Working? Selection, Treatment, and the Market for Remote Work," 2023.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "On her own account: How strengthening women's financial control impacts labor supply and gender norms," *American Economic Review*, 2021, *111* (7), 2342–2375.

Filer, Randall K, "Male-female wage differences: The importance of compensating differentials," *ILR Review*, 1985, *38* (3), 426–437.

Fletcher, Erin, Rohini Pande, and Charity Maria Troyer Moore, "Women and work in India: Descriptive evidence and a review of potential policies," 2017.

Goldin, Claudia, "A grand gender convergence: Its last chapter," *American Economic Review*, 2014, *104* (4), 1091–1119.

_ , *Career & Family: Women's Century-Long Journey toward Equity*, Princeton University Press, 2021.

_ **and Lawrence F Katz**, "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation," *Journal of Labor Economics*, 2016, *34* (3), 705–746.

Gulesci, Selim, Sam Jindani, Eliana La Ferrara, David Smerdon, Munshi Sulaiman, and H Young, "A Stepping Stone Approach to Norm Transitions," 2023.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The allocation of talent and us economic growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Jalota, Suhani and Lisa Ho, "What Works For Her? How Work-from-Home Digital Jobs Affect Female Labor Force Participation," *Job Market Paper*, 2023.

Jayachandran, Seema, "Social norms as a barrier to women's employment in developing countries," *IMF Economic Review*, 2021, *69* (3), 576–595.

Karlan, Dean and Jonathan Zinman, "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment," *Econometrica*, 2009, *77* (6), 1993–2008.

Khanna, Madhulika and Divya Pandey, "Reinforcing gender norms or easing housework burdens? The role of mothers-in-law in determining women's labor force participation," 2021.

Kohlberg, Lawrence, "Moral stages and moralization: The cognitive-development approach," *Moral development and behavior: Theory research and social issues*, 1976, pp. 31–53.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Lowe, Matt and Madeline McKelway, "Coupling labor supply decisions: An experiment in india," 2021.

Markus, Hazel and Paula Nurius, "Possible selves.," *American psychologist*, 1986, *41* (9), 954.

Mas, Alexandre and Amanda Pallais, "Valuing alternative work arrangements," *American Economic Review*, 2017, *107* (12), 3722–3759.

McKelway, Maddie, "Information, Norms, and Female Employment: An Experiment in India," 2023.

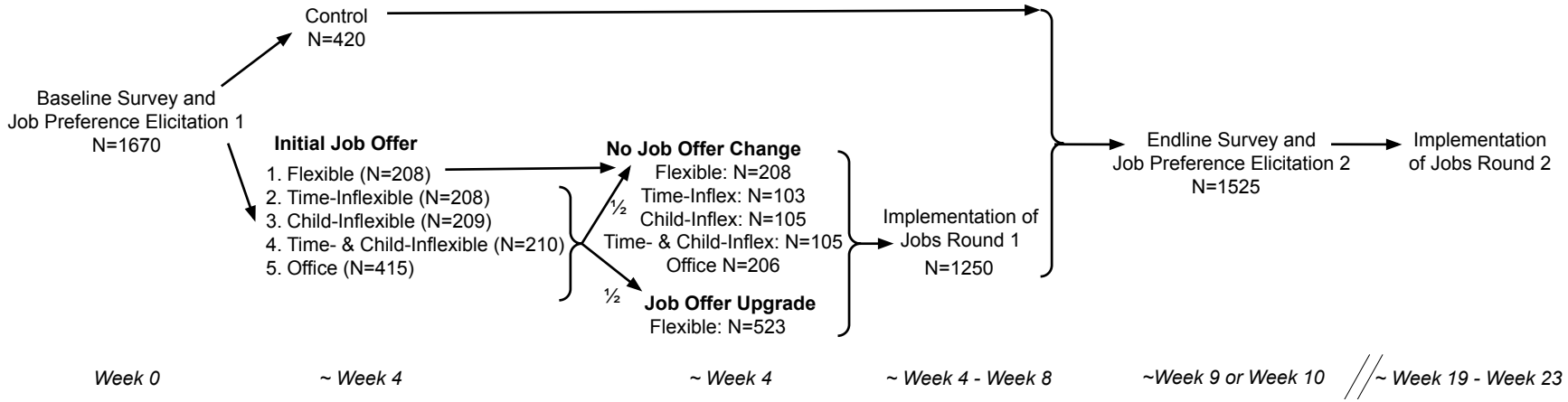
McKelway, Madeline, "Women's employment in India: Intra-household and intra-personal constraints," *Working paper*, 2021.

Oh, Suanna, "Does identity affect labor supply?," *American Economic Review*, 2023, *113* (8), 2055–2083.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

Figures

Figure 1: Experimental Design

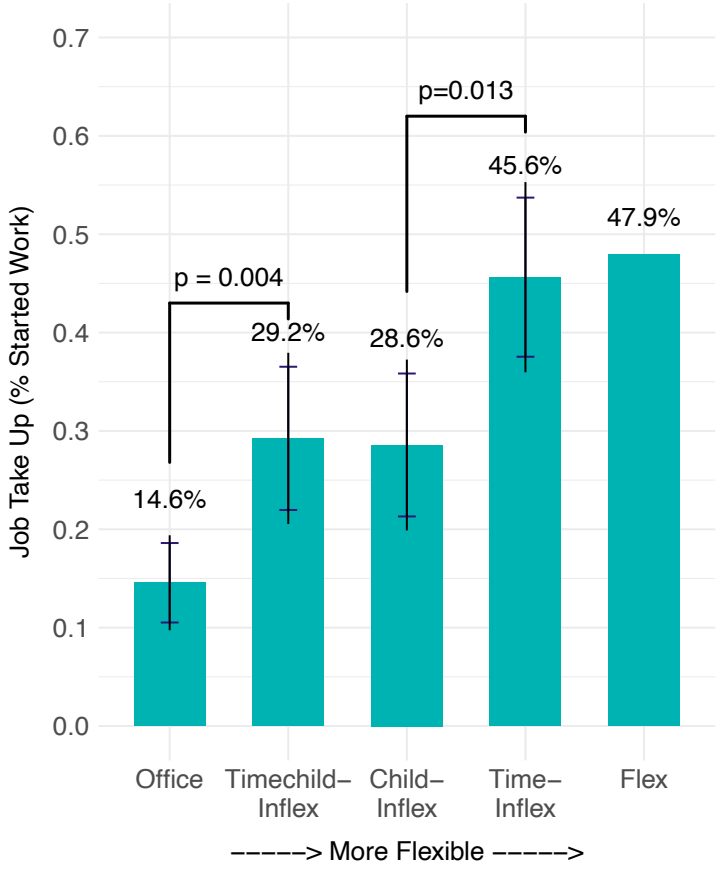


Notes: This figure visualizes the experimental design and timeline.

- Eligible households complete a baseline survey, with one part for women and one optional part for children aged 8-18. The baseline survey for women includes modules about (i) demographics, household composition, and previous work experience, (ii) technology use, (iii) time use, (iv) gender attitudes, (v) agency, (vi) psychological wellbeing, (vii) bank use, and (viii) social contacts. As part of the baseline survey, women also complete a job preferences elicitation which involves stating whether or not they would accept each of the five work arrangements offered as part of the study. The baseline survey for children includes modules about (i) aspirations, (ii) gender attitudes, and (iii) their role in home production.
- 1,670 households are randomized into receiving a job offer or to the control group. The jobs vary along three dimensions: (1) the ability to flexibly choose work hours, (2) the ability to multitask work with childcare, and (3) the ability to work from home. Time-flexibility and childcare-flexibility are cross-randomized, resulting in five job groups.
- After deciding whether or not to accept the job offer, half of the participants who initially received an inflexible job are randomly selected for an upgrade to the most flexible job. This surprise upgrade allows us to separately measure selection into flexible work arrangements (characterizing the "flexibility compliers") and to estimate the treatment effects of the flexible work arrangements on job performance, mirroring the design in Karlan and Zinman (2009). After this final job offer, participants start their part-time, month-long job that consists of microtasks that can be done on a smartphone. The purpose of the microtasks is to build datasets to train speech recognition algorithms.
- Within two weeks of job completion, participating women and children complete an endline survey. The children's survey includes the same modules as the baseline survey, with some questions modified. The endline survey for women includes modules on (i) household members' labor supply, (ii) gender attitudes, (iii) agency, (iv) psychological wellbeing, and, if the woman participated in the intervention, (v) her experience with the job.
- As part of the endline survey, women also complete another job preferences elicitation that involves making 7 incentivized choices between jobs and gifts. We use the strategy method to incentivize the choices, randomly selecting one of the decisions to be implemented as "Jobs Round 2." This second round of jobs includes digital and non-digital job options, and varies in flexibility along the same dimensions as the initial intervention (work hours, multitasking work with childcare, and working from home). The digital jobs are the same as in the initial intervention (contributing to speech datasets), and the non-digital jobs involve sewing masks and making jewellery. In order to estimate a real-stakes treatment effect on interest in future work, jobs in the second round are fully implemented for the same duration as the initial intervention jobs.

Figure 2: Impact of flexible work arrangements on take up of jobs

(a) Job Take Up Rates



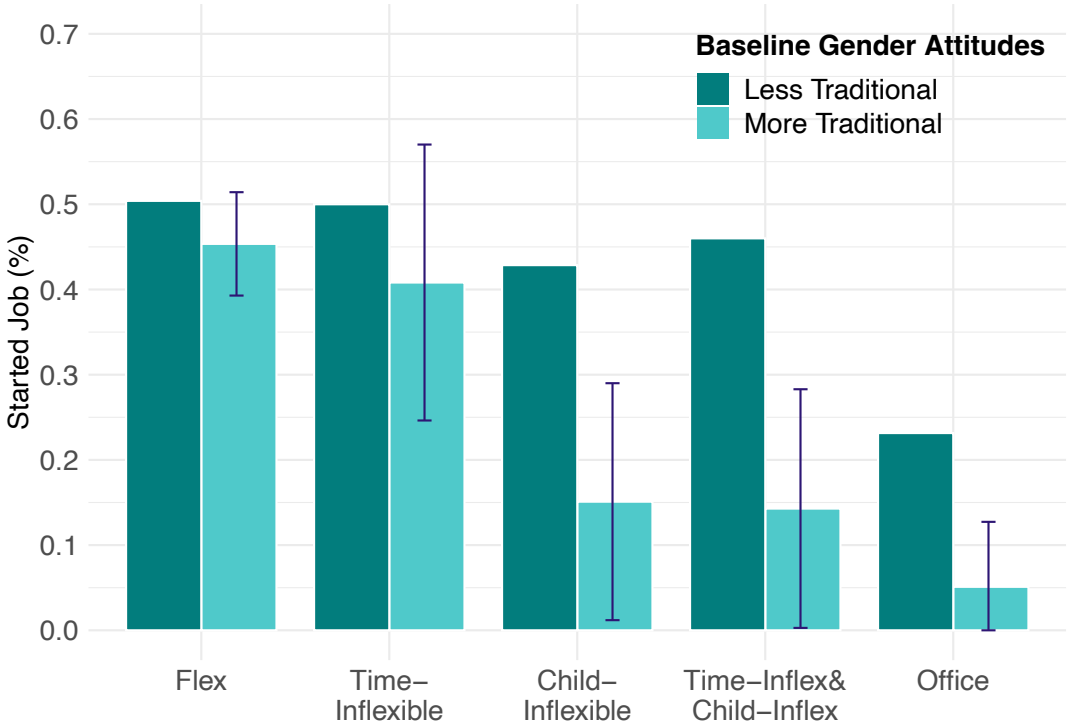
(b) Job Descriptions

Final Work Arrangement (N=1250)	Choose Hours	Multitasking (Childcare)	Work from Home
1. Most Flexible (N=731)	Yes	Yes	Yes
2. Time-Inflexible (N=103)	No	Yes	Yes
3. Child-Inflexible (N=105)	Yes	No	Yes
4. Time-Inflexible & Child-Inflexible (N=105)	No	No	Yes
5. Office (N=206)	(Yes)	No	No

Notes: This figure plots the take up rate for each of the five jobs during the initial intervention (Jobs Round 1).

- Panel A plots job take up for the 1,250 treatment group participants, each of whom receives one job offer. Take up is measured as a binary variable equal to one if the participant starts work (i.e. submits completed tasks to the employer). The whiskers indicate 90% and 95% confidence intervals from a regression of job take up on dummy variables for each of the four jobs other than "Flex," which is the most flexible work arrangement. The estimates and standard errors for these regressions, along with pairwise tests of equality between job take up rates, are presented in column 3 of Table 2.
- The table in Panel B describes how the five jobs sequentially turn on the ability to (1) choose work hours flexibly, (2) multitask work with childcare, and (3) work from home.

Figure 3: Heterogeneity in Job Take Up by Baseline Gender Attitudes

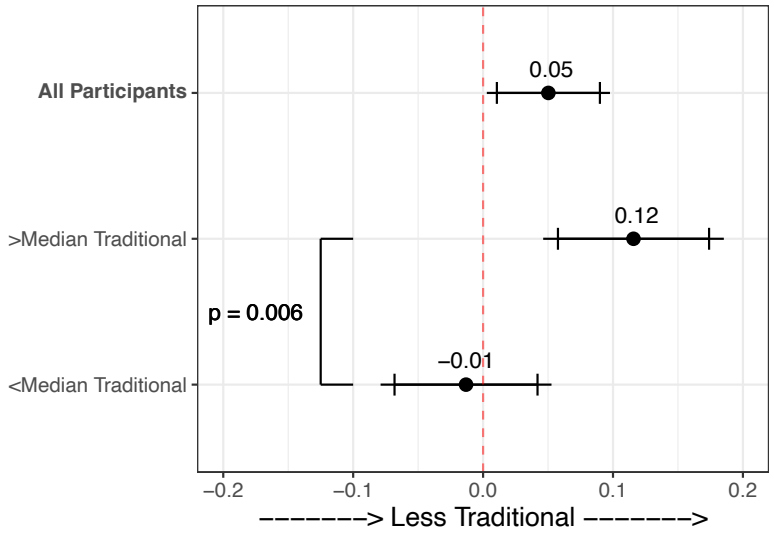


Notes: This figure shows the differential job take up rates (starting work) by baseline household gender attitudes for the 1,250 treatment group participants.

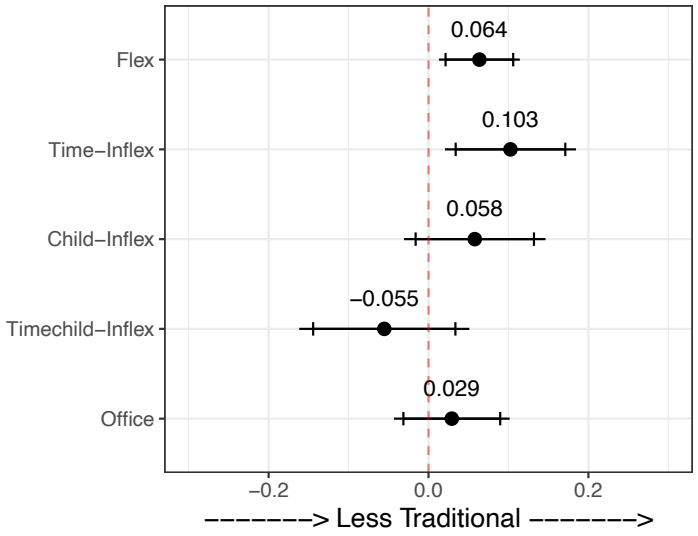
- Take up is measured as a binary variable equal to one if the participant started work (i.e. submitted completed tasks) after the job offer.
- To compare job take up in each category between more traditional and less traditional participants, we regress take up on a binary variable equal to one if the participant's baseline gender attitudes were more traditional than the median participant. Gender attitudes are computed as a weighted average of 16 questions, in which the weights take into account the covariance structure of the components (Anderson, 2008).
- To see results of regressions that estimate the differential importance of job flexibility by baseline gender attitudes, see Table 3. The results presented in the table come from regressions that also control for household characteristics such as income, age, household composition, education, and religion.
- Confidence intervals at the 90% level are shown.

Figure 4: Treatment Effect on Women's Own Gender Attitudes

(a) Heterogeneity by Baseline Attitudes



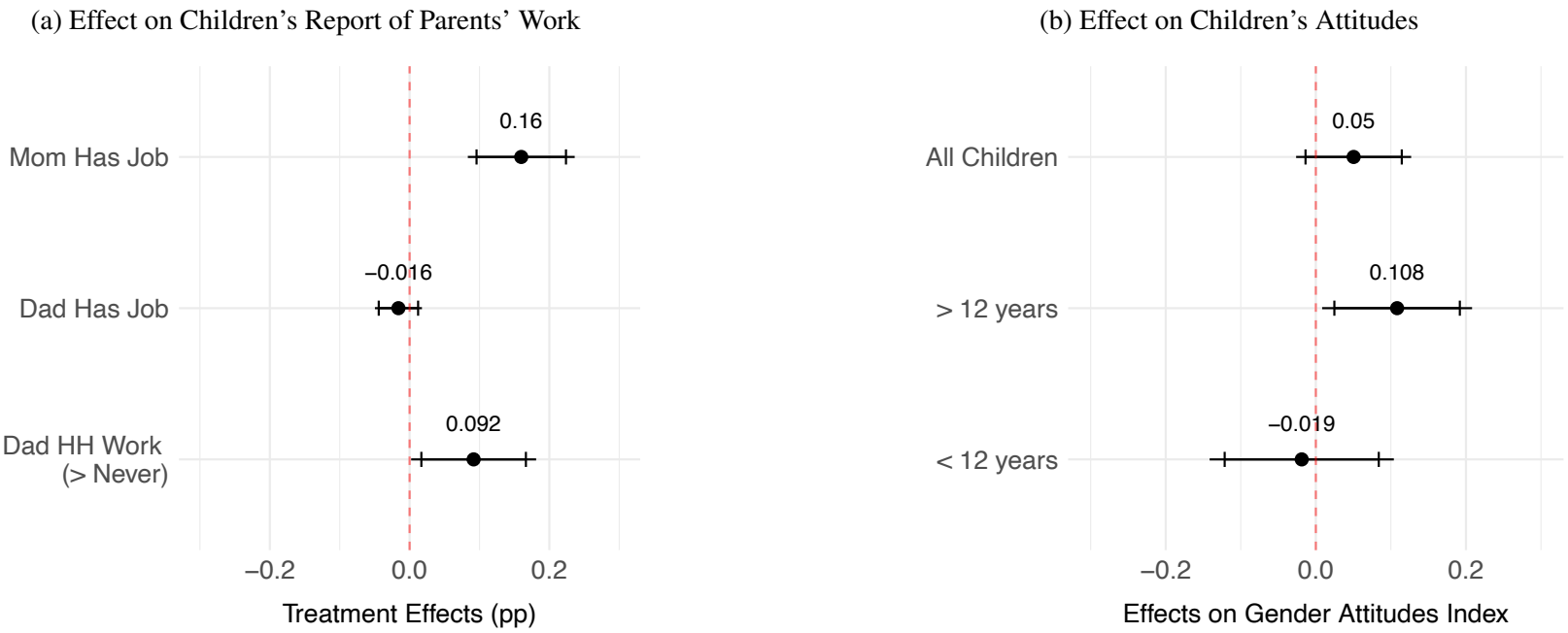
(b) Heterogeneity by Work Arrangement



Notes: This figure reports intent-to-treat effects on gender attitudes, with heterogeneity by baseline attitudes (Panel A) and randomly assigned work arrangement (Panel B).

- In both Panels A and B, participants' scores on the endline gender attitudes index are regressed on treatment assignment. In Panel A, the regression is first on all participants (top), and then separately on participants with pre-intervention attitudes more traditional than the median (middle) and less traditional than the median (bottom). The $p-$values correspond to a test of whether the treatment effect on gender attitudes is equal in the two subsamples. In Panel B, the effect on each participants randomly assigned to each work arrangement is estimated separately by work arrangement. All specifications include lasso-selected controls and strata fixed effects.
- The gender attitudes index is computed as a weighted average of questions from the baseline survey or the endline survey, in which the weights take into account the covariance structure of the components (as in Anderson, 2008).
- In Panel A, participants are categorized according to whether their baseline gender attitudes are more (or less) traditional than the median study participant.
- In Panel B, participants are categorized according to the work arrangement they were randomly assigned to.
- Standard errors are heteroskedasticity robust. Estimates are plotted along with corresponding 90% and 95% confidence intervals.

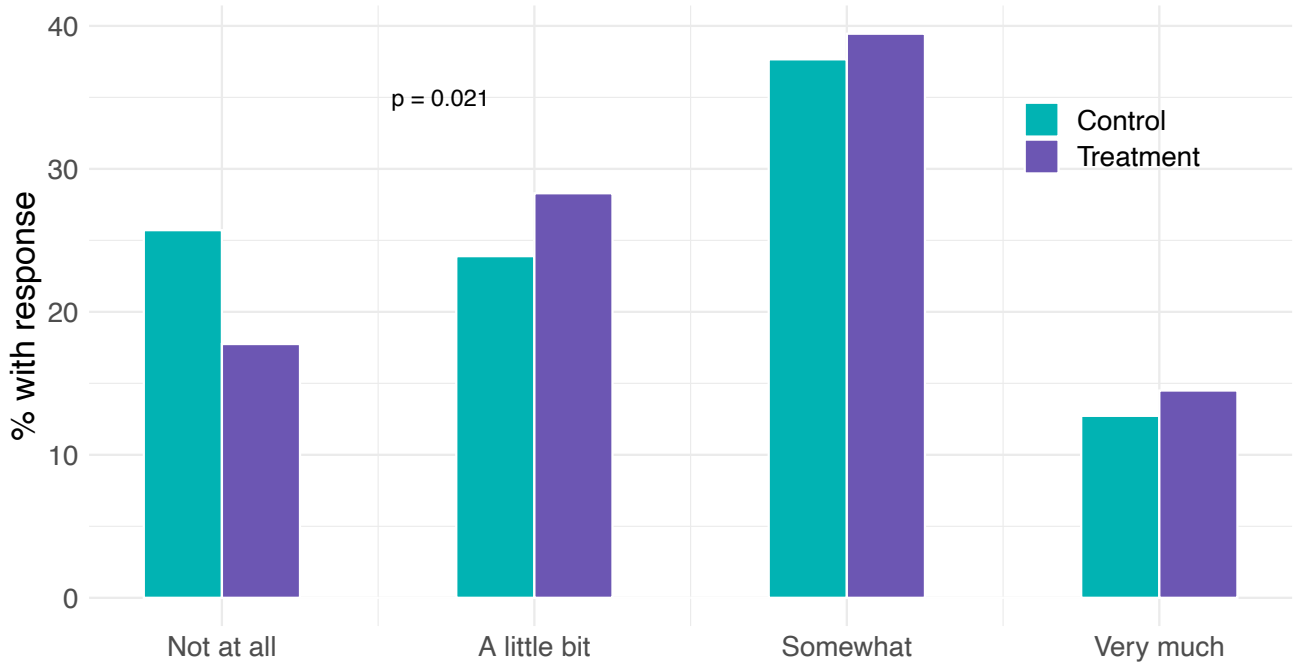
Figure 5: Treatment Effects on Children



Notes: This figure reports intent-to-treat effects on children, including controls for the baseline survey measure of the outcome when possible.

- Panel A reports effects on whether or not children say that their mother had a job in the last month during the intervention (top), whether or not children say that their father had a job in the last month during the intervention (middle), and a binary variable for whether their father ever helped with childcare, cooking, or cleaning in the last month (bottom).
- Panel B reports effects on children's gender attitudes, first for all children pooled together (top), and then for children older than the median age of 12 (middle), and then for younger children (bottom). The gender attitudes index is computed as a weighted average of questions from the endline survey, in which the weights take into account the covariance structure of the components (as in Anderson, 2008).
- Standard errors are heteroskedasticity robust. Estimates are plotted along with corresponding 90% and 95% confidence intervals.

Figure 6: Perception of Whether Talents and Abilities Are Put to Good Use

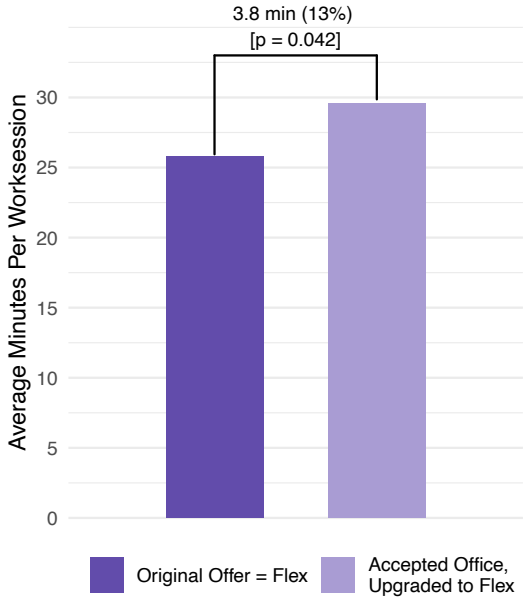


Notes: This figure plots participants' responses to the question, "Do you feel that your full potential and talents are put to good use?" during the endline survey, separating treatment and control participants.

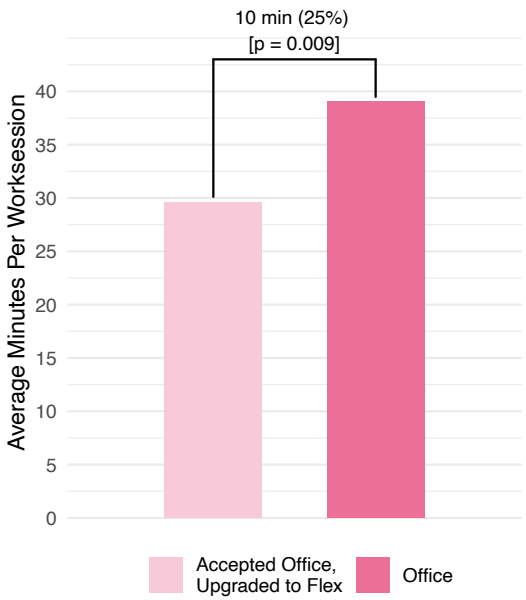
- The green bars represent the fraction of control group participants with each response, while the blue bars represent the fraction of treatment group participants with each response. All treatment groups are pooled in this analysis.
- The reported p-value comes from a regression which codes participants' responses numerically, with "not at all" as 0, "a little bit" as 1, "somewhat" as 2, and "very much" as 3. The treatment group has a mean score of 1.50, while the control group has a mean score of 1.37.
- This outcome is regressed on treatment assignment, lasso-selected controls, and strata fixed effects. Standard errors are heteroskedasticity-robust.

Figure 7: Flow Effects

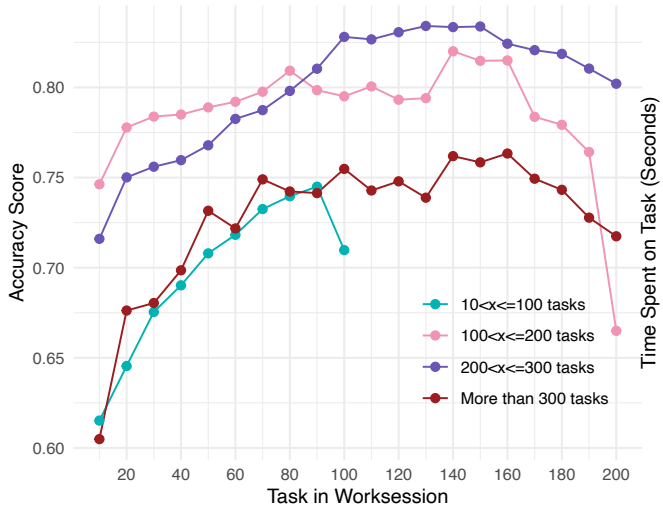
(a) Selection: Worksession Length



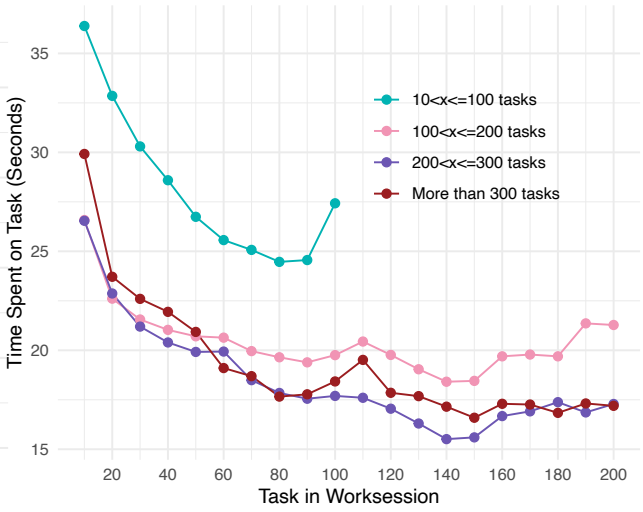
(b) Treatment Effect: Worksession Length



(c) Task accuracy, by task within worksesion



(d) Time spent on task, by task within worksession



Notes: This figure offers evidence for worksession interruptions as a mechanism driving effects on job performance.

- Panel A reports the selection effects of work-from-home on worksession length (minutes elapsed). A worksession is defined as a continuous stretch of work time during which no more than 10 minutes elapse between consecutive tasks. Panel A compares workers who accept the office job and are randomly selected for an upgrade to the most flexible job (lighter bar) with workers who were initially assigned to the most flexible job (darker bar).
- Panel B reports the treatment effects of work-from-home on worksession length (minutes elapsed). The subfigure compares workers who accept an office job and are randomly selected for an upgrade to the most flexible job (lighter bar) with workers who accept an office job and are not randomly selected for an upgrade (darker bar).
- Panels C and D describe how the two key inputs into productivity (average time spent on a task, and task accuracy) change over the course of a worksession. In order to capture "flow effects" rather than selection into longer versus shorter worksessions, worksessions are first grouped according to their number of tasks.

Tables

Table 1: Balance Table (Participants Who Completed the Endline Survey)

	All Participants	Control	Treatment	Pairwise t-test
Endline Complete (=1)	0.913	0.919	0.911	0.585
Age	29.955	29.640	30.062	0.232
Completed 10th Standard (=1)	0.494	0.526	0.483	0.230
Scheduled Caste/Tribe (=1)	0.395	0.394	0.396	0.995
Hindu (=1)	0.762	0.736	0.771	0.064*
Never Married (=1)	0.069	0.060	0.072	0.332
Number HH Members	4.603	4.687	4.575	0.254
Parent-in-Law in HH (=1)	0.403	0.438	0.391	0.098*
Has Child Under 8 (=1)	0.483	0.474	0.486	0.204
Job Decision Final Say (=Self)	0.362	0.370	0.359	0.729
Has Own Smartphone (=1)	0.725	0.731	0.723	0.642
Gender Attitudes Index	0.029	0.001	0.038	0.102
Agency Index	0.017	-0.005	0.025	0.441
Number of participants	1525	386	1139	
F-test of joint significance (F-stat)				1.251
F-test, number of observations				1525

Notes: This table presents summary statistics and balance checks on participant characteristics. Each row shows the mean for that variable for the entire study population, the control group, and the treatment group. The top row of this table records the endline completion rates for all 1,670 participants who were randomized (420 for the control group and 1,250 for the treatment group). Regressions include strata fixed effects. Significance at the 0.10, 0.05, and 0.01 levels are indicated by *, **, and ***, respectively.

Table 2: Effect of Flexible Job Attributes on Take Up of Work

	Job Take Up		
	Baseline	Job Offer	Start Work
	(1)	(2)	(3)
Time-inflexible	-0.04***	-0.04	-0.02
	(0.01)	(0.04)	(0.05)
	[0.00]	[0.38]	[0.66]
Child-inflexible	-0.27***	-0.17***	-0.19***
	(0.01)	(0.05)	(0.05)
	[0.00]	[0.00]	[0.00]
Time- & Child-inflexible	-0.32***	-0.15***	-0.19***
	(0.01)	(0.05)	(0.05)
	[0.00]	[0.00]	[0.00]
Office	-0.42***	-0.25***	-0.33***
	(0.01)	(0.04)	(0.03)
	[0.00]	[0.00]	[0.00]
Observations	8,290	1,250	1,250
Most Flexible Job Take Up Rate	0.98	0.75	0.48
P-val: equality of coefficients			
Time-inflex == Child-inflex	0.000	0.006	0.010
Time-inflex == Time & Child-inflex	0.000	0.018	0.013
Time-inflex == Office	0.000	0.000	0.000
Child-inflex == Time & Child-inflex	0.000	0.705	0.914
Child-inflex == Office	0.000	0.061	0.006
Time & Child-inflex == Office	0.000	0.020	0.004

Notes: This table presents the impacts of flexible work arrangements on job take up.

- The estimates come from regressions where the outcome variable is take up, and the regressors are dummy variables for each of the four work arrangements that are not the most flexible job ("Flex"). No control variables are included. Take up is measured as a dummy variable equal to 1 if the participant took up the job.
- Each column shows a different definition of job take up. Column (1) measures take up according to whether, on the baseline survey, the participant says she would accept the job if offered it. Column (2) measures take up according to whether the participant says yes when called with a job offer by the jobs team. Column (3) measures take up according to whether the participant actually begins work (i.e. submitted job tasks to the employer for review).
- At baseline, each participant was asked about each of the five work arrangements (in randomized order). At the job offer and starting work stage, each participant had been randomized to one work arrangement. The number of observations in Column (1) is 8290 = 1658 $\times$ 5 because there are 12 respondents who did not respond to the location-inflexible question.
- Standard errors in parentheses $(\cdot)$ are robust to heteroscedasticity and clustered at the participant level for column (1). Brackets $[\cdot]$ report unadjusted p-values, as do stars next to coefficients (* significant at 10%; ** at 5%; *** at 1%).

Table 3: Differential Importance of Flexibility by Gender Attitudes and Young Children

	Baseline (1)	Job Offer (2)	Start Work (3)
Panel A: All Participants			
Flexible Job	0.34***	0.19***	0.34***
	(0.01)	(0.04)	(0.03)
>Median Traditional Gender Atts	-0.06***	-0.12***	-0.23***
	(0.02)	(0.03)	(0.04)
Flexible Job × >Median Traditional Gender Atts	0.04**	0.02	0.17***
	(0.02)	(0.06)	(0.05)
Take Up Rate for < Median Traditional, Inflexible Jobs	0.67	0.60	0.33
P-val: equality of coefficients			
Flexible Job: More Traditional == Less Traditional	0.034	0.026	0.109
Observations	8,290	1,250	1,250
Panel B: Participants without children under 8			
Flexible Job	0.25***	0.17***	0.24***
	(0.01)	(0.06)	(0.05)
>Median Traditional Gender Atts	-0.07***	-0.10**	-0.19***
	(0.02)	(0.05)	(0.06)
Flexible Job × >Median Traditional Gender Atts	0.05***	0.03	0.04
	(0.02)	(0.08)	(0.07)
Take Up Rate for < Median Traditional, Inflexible Jobs	0.77	0.62	0.35
P-val: equality of coefficients			
Flexible Job: More Traditional == Less Traditional	0.238	0.196	0.002
Observations	4,320	649	649
Panel C: Participants with children under 8			
Flexible Job	0.42***	0.21***	0.44***
	(0.02)	(0.06)	(0.04)
>Median Traditional Gender Atts	-0.02	-0.13***	-0.26***
	(0.03)	(0.05)	(0.05)
Flexible Job × >Median Traditional Gender Atts	0.00	0.02	0.30***
	(0.03)	(0.08)	(0.07)
Take Up Rate for < Median Traditional, Inflexible Jobs	0.55	0.58	0.31
P-val: equality of coefficients			
Flexible Job: More Traditional == Less Traditional	0.075	0.065	0.408
Observations	3,970	601	601

Notes: This table presents the heterogeneous importance of job flexibility for take up by baseline gender attitudes and having a young child. Panel A shows heterogeneity by baseline gender attitudes among all participants, while Panels B and C show heterogeneity by baseline gender attitudes among different subgroups (without children under age 8 and with children under 8, respectively).

- Job take up is measured in three ways: in column (1), whether the participant says on the baseline survey that she would accept the job if offered it, in column (2), whether she accepts the job when called by the jobs team, and in column (3), whether she starts work.
- Gender attitudes are measured by a weighted index of 16 questions from the baseline survey, with weights accounting for the covariance structure of the index components (as in Anderson, 2008).
- "Flexible Job" is a dummy variable equal to one for the flexible and time-inflexible jobs, and is equal to zero for all other work arrangements.
- These regressions control for age, marital status, previous employment, completion of 10th standard, living with parents-in-law, religion, number of household members, whether the participant has a child under age eight (in Panel A), smartphone ownership, and region of Kolkata.
- Standard errors in parentheses are robust to heteroskedasticity and clustered at the participant level for column (1). Stars next to coefficients represent unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table 4: Treatment Effect of Jobs Intervention on Gender Attitudes - Heterogeneity by Work Arrangement

	Gender Attitudes Index					
	ITT Estimates			2SLS Estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
A: All Treatments Pooled						
Treatment	0.06**	0.06**	0.05**			
	(0.03)	(0.03)	(0.02)			
Started Work				0.15**	0.15**	0.12**
				(0.06)	(0.06)	(0.06)
B: Heterogeneity by Work Arrangement						
Arrangement = Flex or Time-Inflex	0.08***	0.08***	0.07***			
	(0.03)	(0.03)	(0.03)			
Arrangement = Child- or Child- & Time-Inflex	0.00	0.00	0.00			
	(0.04)	(0.04)	(0.04)			
Arrangement = Office	0.05	0.05	0.03			
	(0.04)	(0.04)	(0.04)			
Started Flex or Time-Inflex				0.16***	0.16***	0.13***
				(0.05)	(0.05)	(0.05)
Started Child- or Child- & Time-Inflex				0.00	-0.01	-0.01
				(0.13)	(0.13)	(0.12)
Started Office				0.34	0.30	0.18
				(0.25)	(0.25)	(0.23)
Strata Fixed Effects		Yes	Yes		Yes	Yes
Lasso Selected Controls			Yes			Yes
Observations	1,525	1,525	1,525	1,525	1,525	1,525
Control Mean	0.01	0.01	0.01	0.01	0.01	0.01

Notes: This table presents intent-to-treat (ITT) and two-stage-least-squares (2SLS) results for the effect of the intervention on gender attitudes. Panel A presents results for the effects when pooling all the jobs groups together, while Panel B examines heterogeneity by work arrangement.

- The gender attitudes index is computed as a standardized, weighted average of questions from the endline survey, in which the weights take into account the covariance structure of the components (Anderson, 2008).
- The dependent variable is gender attitudes measured on the endline survey, which took place within two weeks of the end of the job treatments.
- Heteroskedasticity- robust standard errors are reported in parentheses. Stars next to coefficients represent unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table 5: Treatment Effect of Jobs on Gender Attitudes - Heterogeneity by Baseline Attitudes

	Gender Attitudes Index					
	ITT Estimates			2SLS Estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.10***	0.10***	0.11***			
	(0.04)	(0.04)	(0.04)			
	[0.008]	[0.006]	[0.001]			
<Traditional at Baseline	0.31***	0.30***	0.06	0.31***	0.30***	0.07
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
	[0.000]	[0.000]	[0.215]	[0.000]	[0.000]	[0.153]
Treatment × <Traditional at Baseline	-0.10**	-0.11**	-0.13***			
	(0.05)	(0.05)	(0.05)			
	[0.041]	[0.025]	[0.006]			
Started Work				0.26***	0.27***	0.31***
				(0.10)	(0.10)	(0.10)
				[0.008]	[0.006]	[0.001]
Started Work × <Traditional at Baseline				-0.27**	-0.29**	-0.34***
				(0.12)	(0.12)	(0.12)
				[0.025]	[0.016]	[0.004]
Strata Fixed Effects		Yes	Yes		Yes	Yes
Lasso-Selected Controls			Yes			Yes
Observations	1,525	1,525	1,525	1,525	1,525	1,525
Control Mean (Gender Attitudes, Endline)	0.01	0.01	0.01	0.01	0.01	0.01
P-val: equality of coefficients						
Treatment Effect (< Traditional) == 0	0.880	0.767	0.557	0.880	0.784	0.566

Notes: This table tests for the impact of the job treatments on gender attitudes, examining heterogeneity by baseline attitudes.

- In this analysis, all treatment groups (work arrangements) are pooled together.
- The dependent variable is gender attitudes measured on the endline survey, which took place within two weeks of the end of the job treatments. The gender attitudes index is computed as a standardized, weighted average of questions from the baseline survey or the endline survey, in which the weights take into account the covariance structure of the components (Anderson, 2008).
- Participants are categorized as "less traditional" if their baseline gender attitudes index value is lower than the median participant.
- Standard errors in parentheses $(\cdot)$ are robust to heteroskedasticity. Brackets $[\cdot]$ report unadjusted p-values, as do stars next to coefficients (* significant at 10%; ** at 5%; *** at 1%).

Table 6: Treatment Effects on Children's Gender Attitudes

	Mom has job	Dad has job	Children's Gender Attitudes		
	(1)	(2)	(3)	(4)	(5)
Panel A: All Children					
Treatment	0.16***	-0.02	0.06	0.07*	0.04
	(0.04)	(0.02)	(0.04)	(0.04)	(0.04)
	[0.000]	[0.352]	[0.160]	[0.080]	[0.198]
Control Mean	0.19	0.97	0.00	0.00	0.00
Panel B: Heterogeneity by Age					
Treatment	0.13**	-0.02	0.10*	0.12**	0.11**
	(0.05)	(0.03)	(0.06)	(0.05)	(0.05)
	[0.011]	[0.368]	[0.076]	[0.029]	[0.034]
Treatment × Age < 12	0.03	0.01	-0.08	-0.10	-0.13
	(0.08)	(0.03)	(0.09)	(0.09)	(0.08)
	[0.707]	[0.821]	[0.346]	[0.243]	[0.120]
Age < 12	-0.03	0.03	-0.01	0.02	0.07
	(0.06)	(0.03)	(0.08)	(0.07)	(0.07)
	[0.667]	[0.231]	[0.856]	[0.795]	[0.307]
Control Mean, Age > 12	0.20	0.96	0.01	0.01	0.01
Panel C: Heterogeneity by Gender					
Treatment	0.19***	0.00	0.06	0.06	0.05
	(0.05)	(0.03)	(0.07)	(0.06)	(0.06)
	[0.000]	[0.895]	[0.342]	[0.340]	[0.046]
Treatment × Male	-0.09	-0.05	-0.02	0.00	-0.01
	(0.07)	(0.03)	(0.09)	(0.08)	(0.08)
	[0.244]	[0.187]	[0.803]	[0.973]	[0.935]
Male	0.16***	0.02	-0.14*	-0.17**	-0.13*
	(0.06)	(0.03)	(0.08)	(0.07)	(0.07)
	[0.008]	[0.505]	[0.070]	[0.025]	[0.062]
Control Mean, Female	0.13	0.96	0.08	0.08	0.08
Observations	601	597	606	606	606
Strata Fixed Effects	Yes	Yes		Yes	Yes
Children's Baseline Attitudes					Yes
Mom Baseline Job Status	Yes				
Dad Baseline Job Status		Yes			

Notes: This table presents results on whether children notice the intervention and subsequent effects on children's gender attitudes.

- In addition to the main effect (Panel A), we present heterogeneous treatment effects by children's age compared to the median age of 12 (Panel B) and child gender (Panel C). In this analysis, all treatment groups (work arrangements) are pooled together.
- Columns (1) and (2) test whether children notice that their mothers are working. In column (1), the outcome variable is a dummy equal to one if children say that their mother had a job in the last month. Column (2) asks the same question, but about their father, as a placebo check. Children who say that they do not know whether or not their mother (or father) is working are dropped. Both of these regressions control for the children's baseline report of their parents' labor supply.
- In columns (3)-(5), the outcome variable is children's gender attitudes. Gender attitudes are computed as a standardized, weighted average of questions from the children's surveys, in which the weights take into account the covariance structure of the components (as in Anderson, 2008).
- Standard errors in parentheses $(\cdot)$ are robust to heteroskedasticity. Brackets $[\cdot]$ report unadjusted p-values, as do stars next to coefficients (* significant at 10%; ** at 5%; *** at 1%).

Table 7: "Gateway jobs": Effect of experience with more flexible job on less flexible job take up

	Started Work in R2 (Round 2)			
	(1)	(2)	(3)	(4)
R1 More Flexible Than R2	0.05*	0.05*	0.07**	
	(0.03)	(0.03)	(0.03)	
R1 Less or Equally Flexible Than R2	-0.02	-0.02	-0.02	
	(0.04)	(0.05)	(0.04)	
Previously Worked = 1			0.04	
			(0.04)	
R1 More Flexible Than R2 × Previously Worked = 1			-0.05	
			(0.06)	
R1 Less or Equally Flexible Than R2 × Previously Worked = 1			-0.01	
			(0.08)	
R2: Time-Inflexible × R2 Less Flexible Than R1				0.03
				(0.07)
R2: Child-Inflexible × R2 Less Flexible Than R1				0.11
				(0.07)
R2: Time- & Child-Inflexible × R2 Less Flexible Than R1				0.02
				(0.07)
R2: Office × R2 Less Flexible Than R1				0.08**
				(0.04)
Observations	1,524	1,524	1,524	1,524
R2 assignment controls	✓	✓	✓	✓
R1 assignment contols		✓		
Strata fixed effects	✓	✓	✓	✓
Lasso selected controls	✓	✓	✓	✓

Notes: This table presents results on effects of the initial intervention on take up in the second round of jobs.

- Columns (1) and (2) reports estimates from model 3, in which the omitted group is the Round 1 control group. Treatment group participants are then categorized by the relative flexibility of their Round 1 versus Round 2 job assignment. Column (2) adds controls for Round 1 job assignment.
- Column (3) reports heterogeneous treatment effects by previous work experience. The variable "Previously Worked" is equal to 1 if the participant reports that she has previously worked for pay, and equal to 0 if the intervention job is the first time she has ever earned income herself.
- Column (4) reports heterogeneous treatment effects by Round 2 job assignment.
- Standard errors in parentheses $(\cdot)$ are robust to heteroskedasticity. Stars next to coefficients denote significance (* at 10%; ** at 5%; *** at 1%).

A Appendix

Appendix Figures

Figure A.8: Take Up of Jobs: All Work Arrangements, All Take-Up Measures

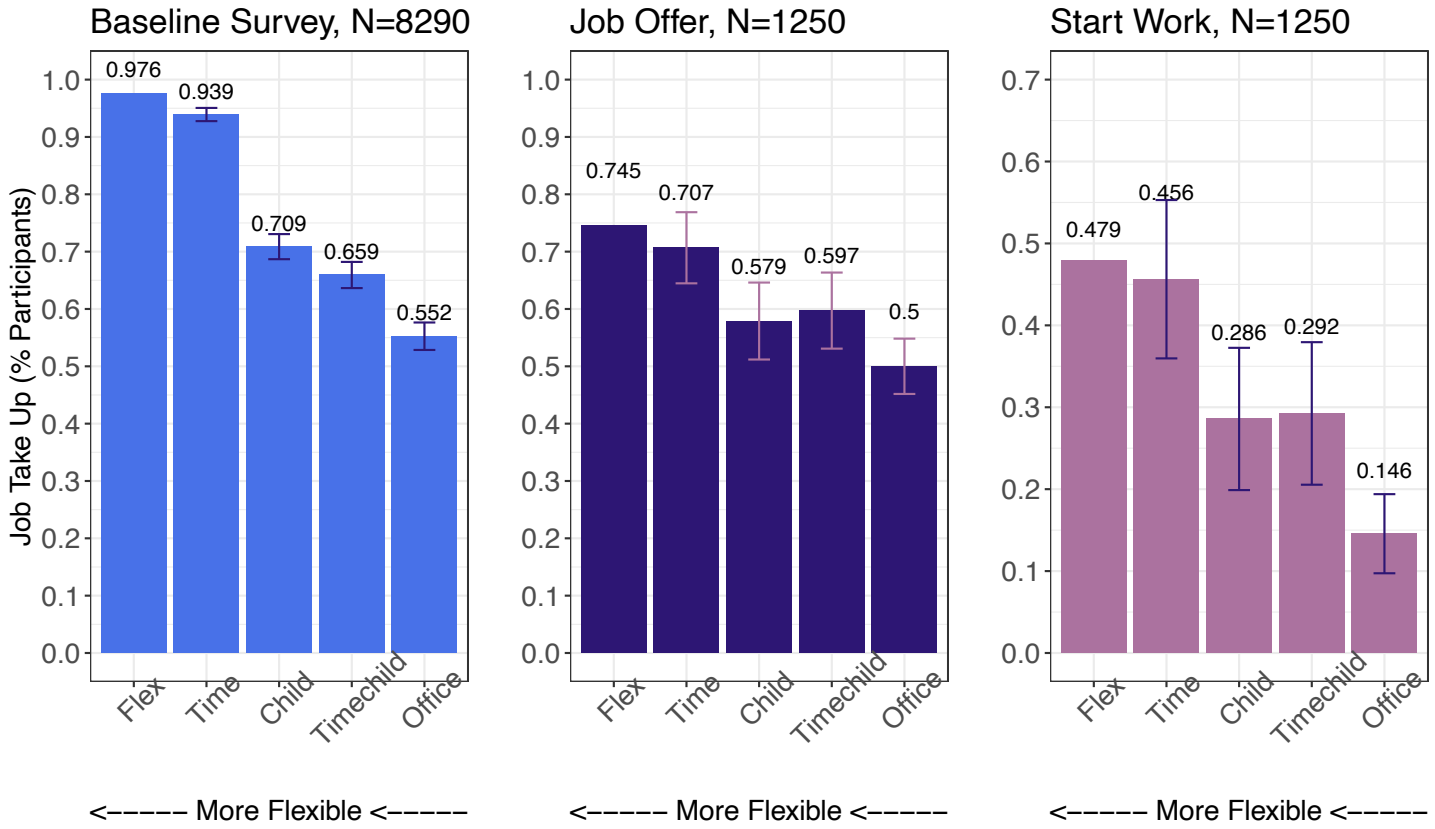
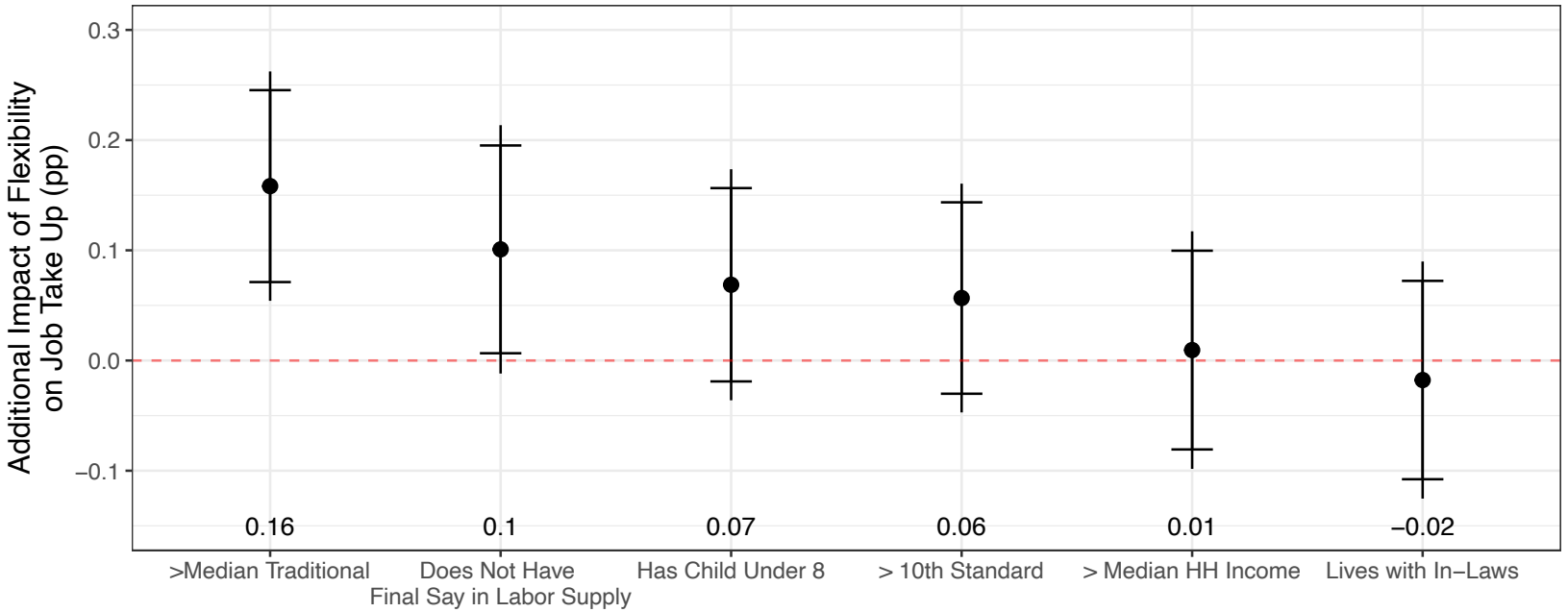


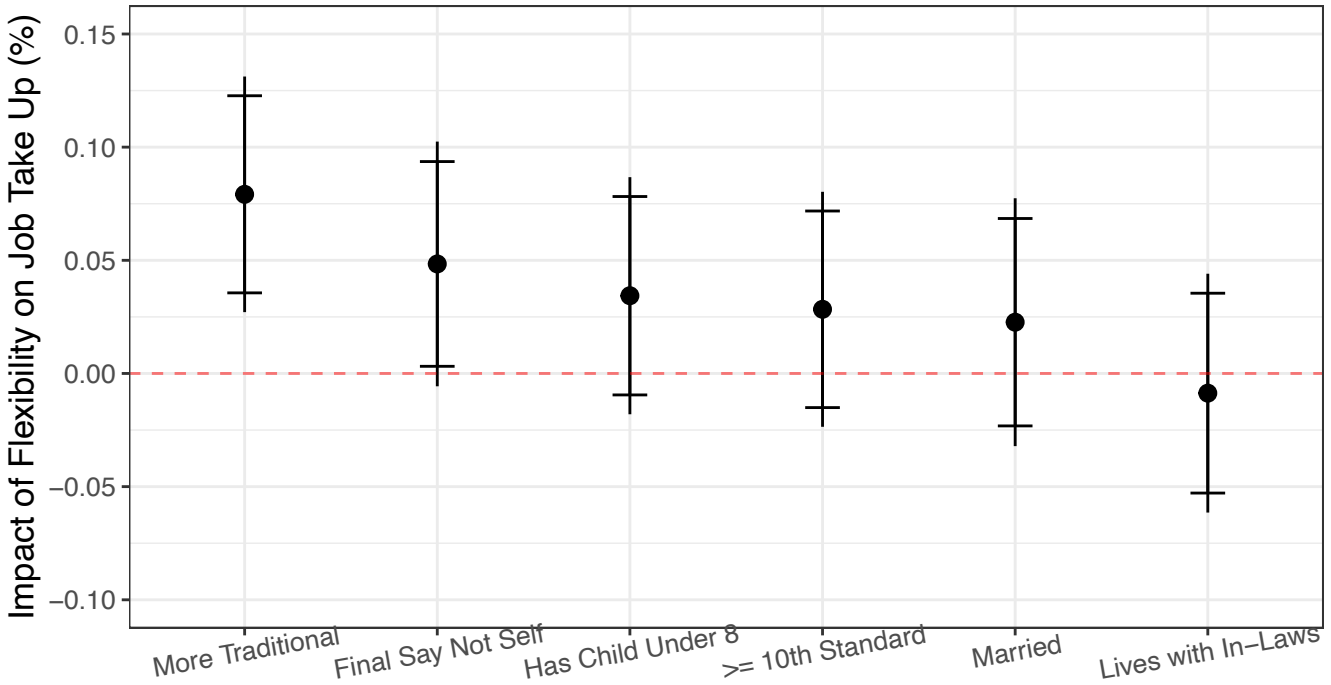
Figure A.9: Heterogeneous Importance of Flexibility for Take Up of Work



Notes: This figure shows how job flexibility affects take up differentially by worker characteristics.

- The plotted coefficients are the β_3 estimates from regressions that take the following form: $y_{ij} = \beta_0 + \beta_1 \text{flexible}_j + \beta_2 \text{characteristic}_i + \beta_3 \text{flexible}_j \times \text{characteristic}_i + \varepsilon_{ij}$, where y_{ij} is a binary variable equal to one if participant i starts job j after being offered it. $flexible_j$ is a dummy variable equal to one for if job j is the most flexible job and equal to zero for all other jobs.
- The β_3 coefficient represents the additional importance of flexibility to job take up for women satisfying a particular characteristic (from left to right: having gender attitudes more traditional than the median participant; not having the final say in one's own labor supply; having a child younger than eight; being educated to at least the 10th grade (standard); having household income greater than the median participating household; living with at least one parent-in-law). All characteristics are binary and are defined to be associated with a greater importance of job flexibility to take up. Gender attitudes are computed as a weighted average of 16 questions on the baseline survey, in which the weights take into account the covariance structure of the components (Anderson, 2008).
- Standard errors are heteroskedasticity-robust. Confidence intervals at the 90% and 95% levels are reported.

Figure A.10: Heterogeneous Importance of Flexibility for Take Up of Work



Notes: Robustness check where "Flexibility" defined as the most flexible job only.

Figure A.11: Gender Attitudes at Baseline and Endline (Histogram)

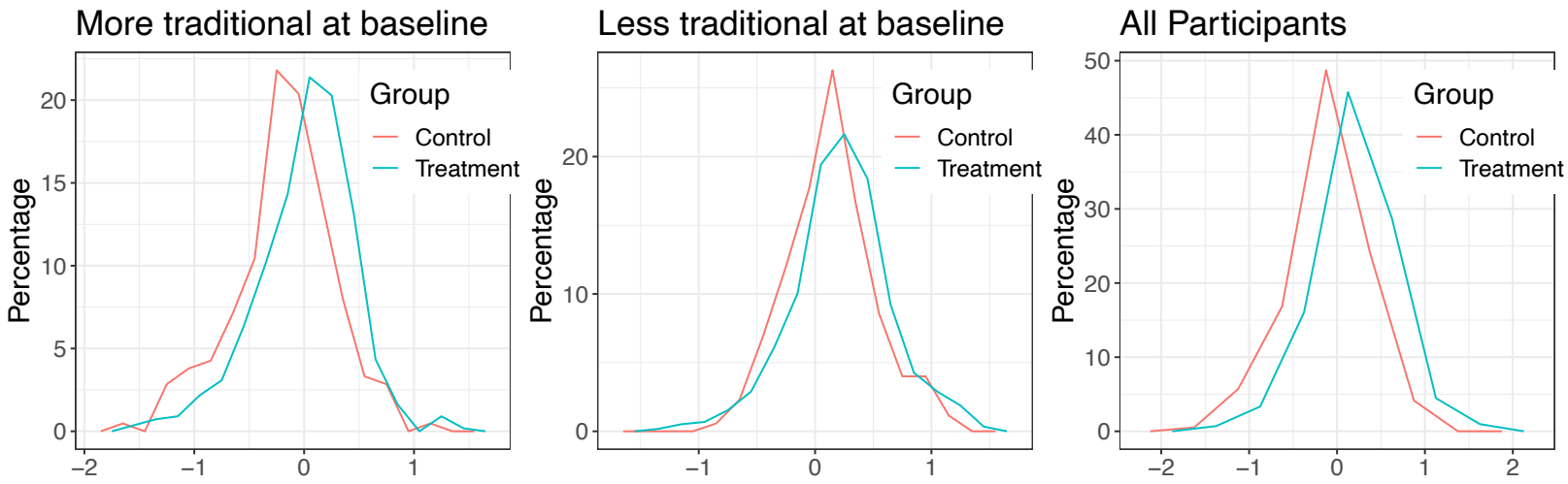
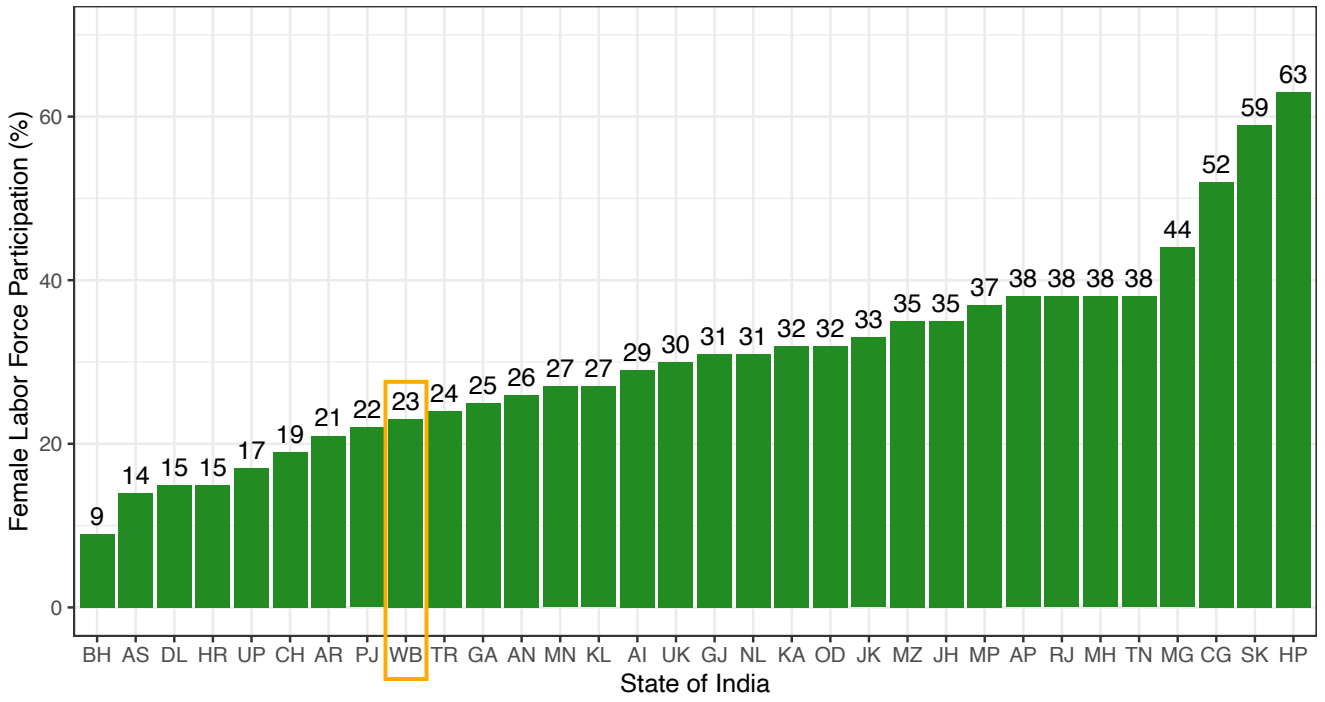


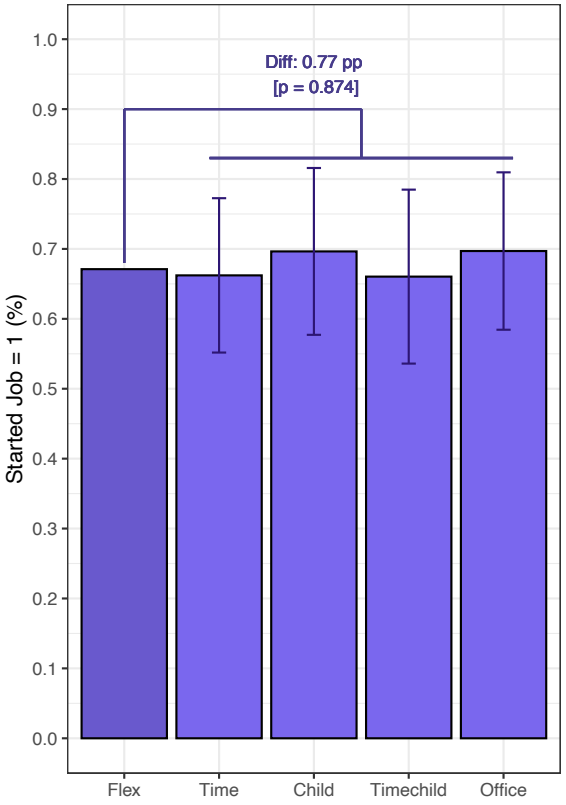
Figure A.12: Female Labor Force Participation Rates (States of India)



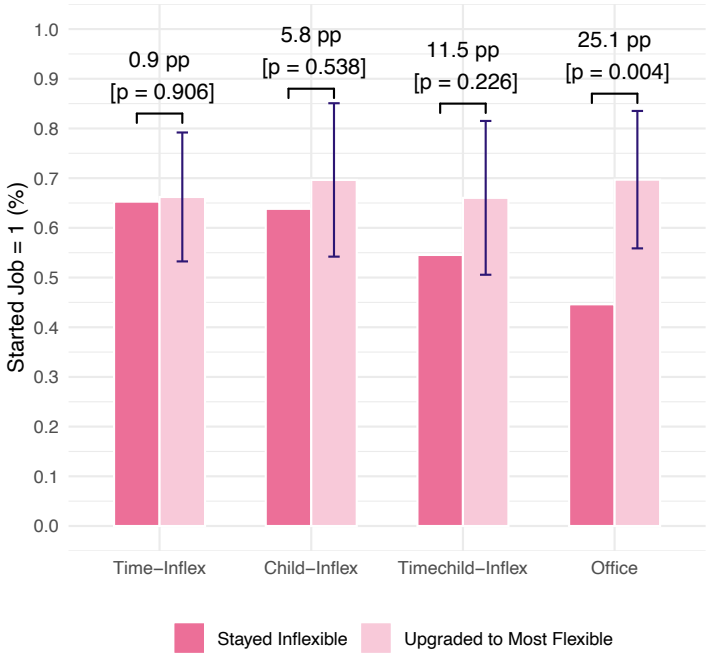
Notes: Data from the Periodic Labor Force Survey 2019-2020. People aged 15 and over included.

Figure A.13: Started Job

(a) Selection



(b) Treatment Effects

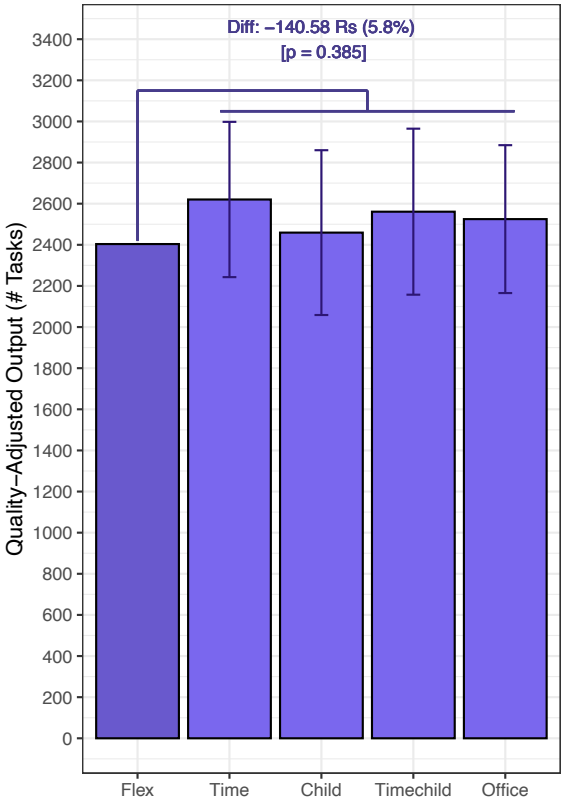


Notes:

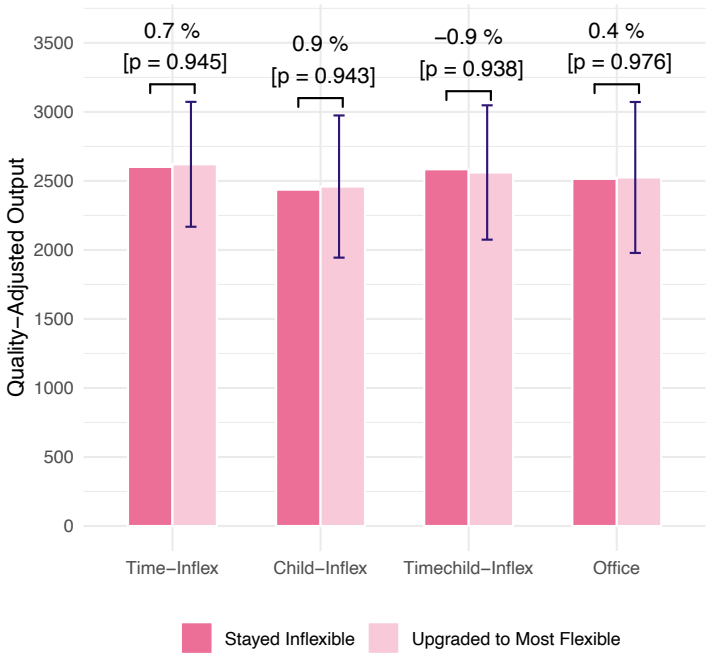
–

Figure A.14: Quality-Adjusted Output (Conditional on Starting Work)

(a) Selection



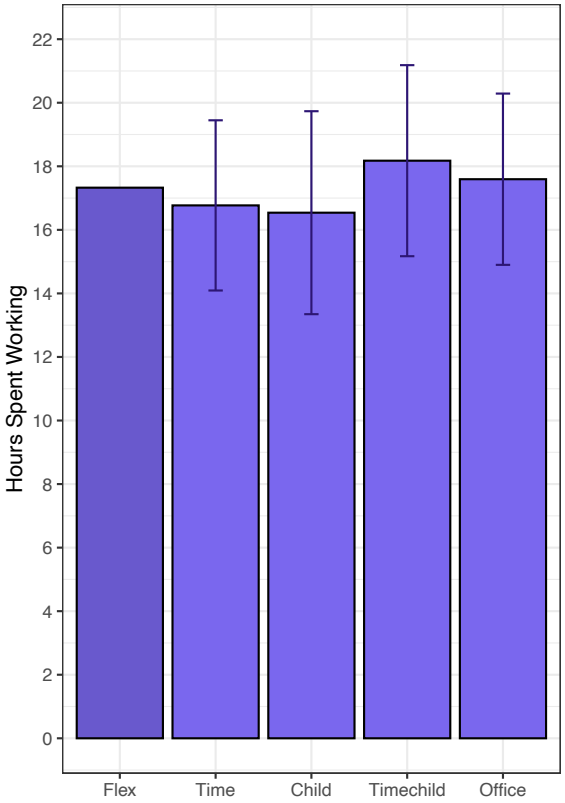
(b) Treatment Effects



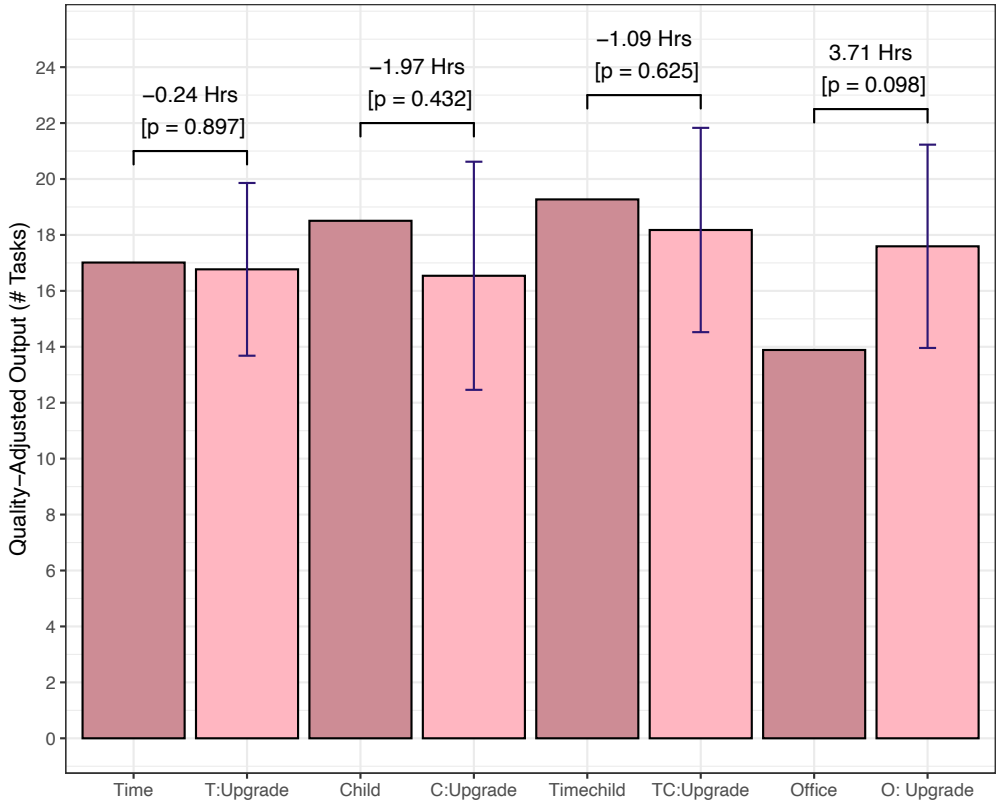
Notes: This figure reports the effects of assignment to flexible and inflexible jobs on gender attitudes (Panel A) and subsequent take up of inflexible jobs during the second round (Panel B). The effects are reported side-by-side to easily allow for comparing effects on attitudes versus job take up.

Figure A.15: Hours Spent (Conditional on Starting Work)

(a) Selection



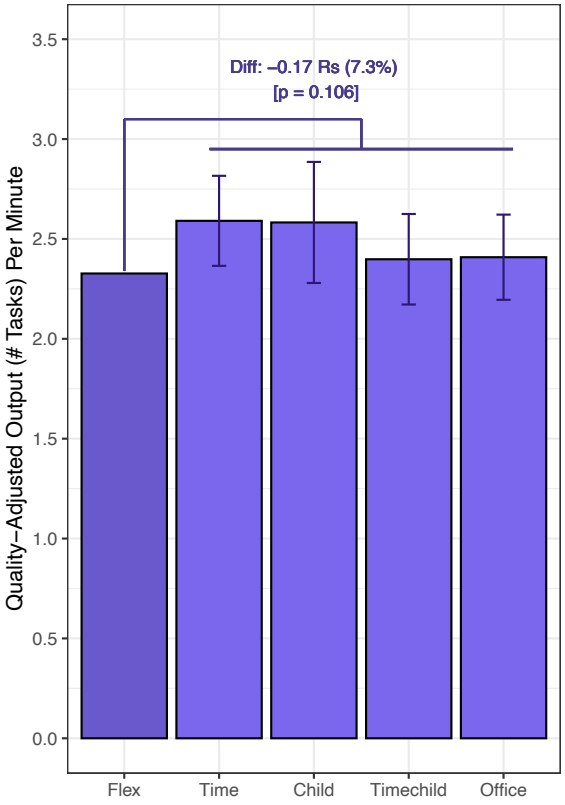
(b) Treatment Effects



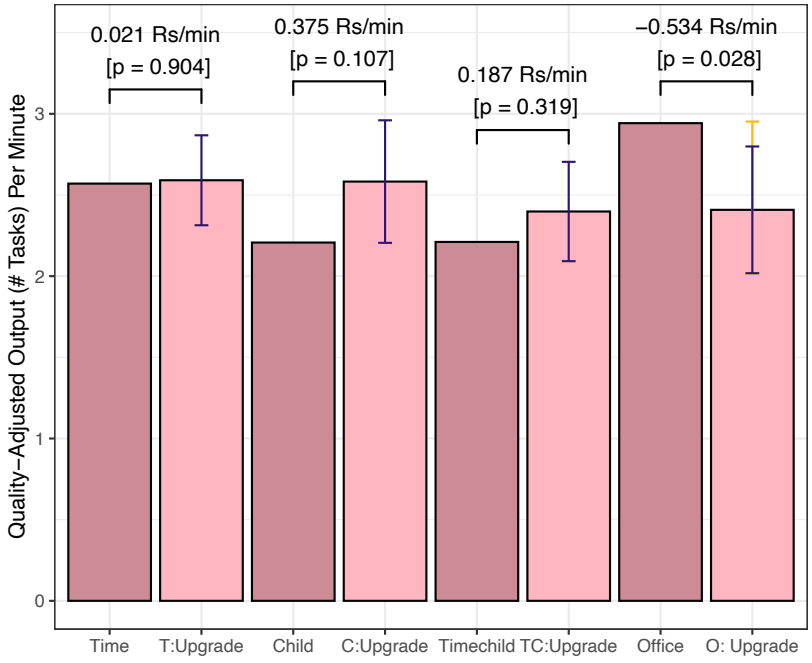
Notes: This figure reports the effects of assignment to flexible and inflexible jobs on gender attitudes (Panel A) and subsequent take up of inflexible jobs during the second round (Panel B). The effects are reported side-by-side to easily allow for comparing effects on attitudes versus job take up.

Figure A.16: Quality Adjusted Output (Earnings) Per Minute

(a) Selection



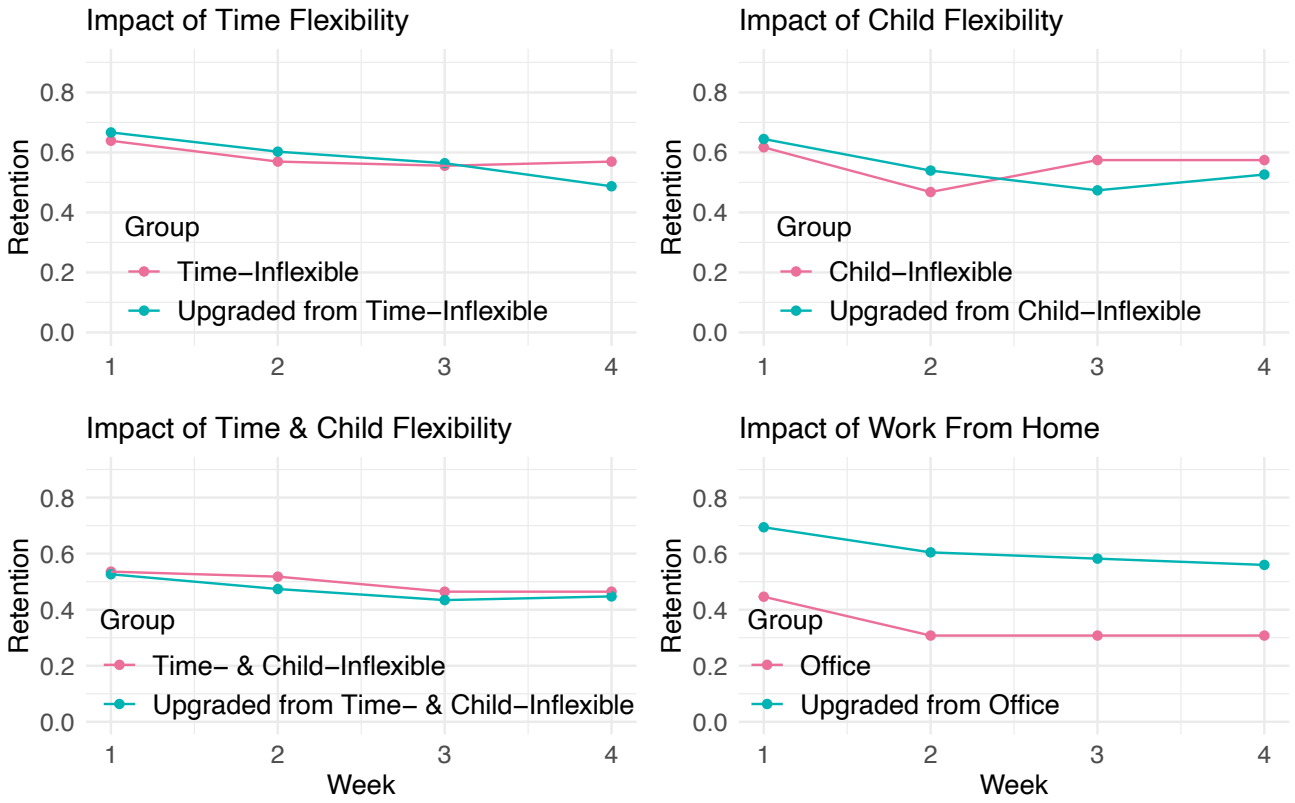
(b) Treatment Effects



Notes:

–

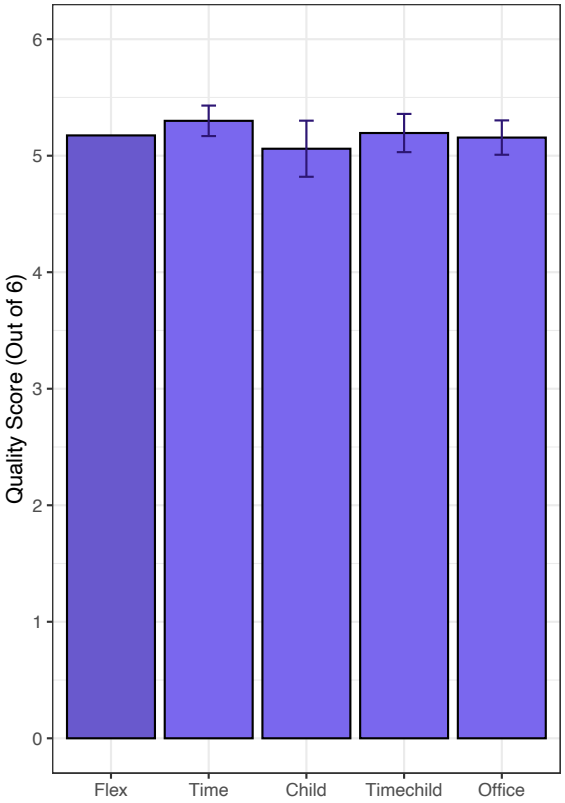
Figure A.17: Impact of job flexibility on worker retention



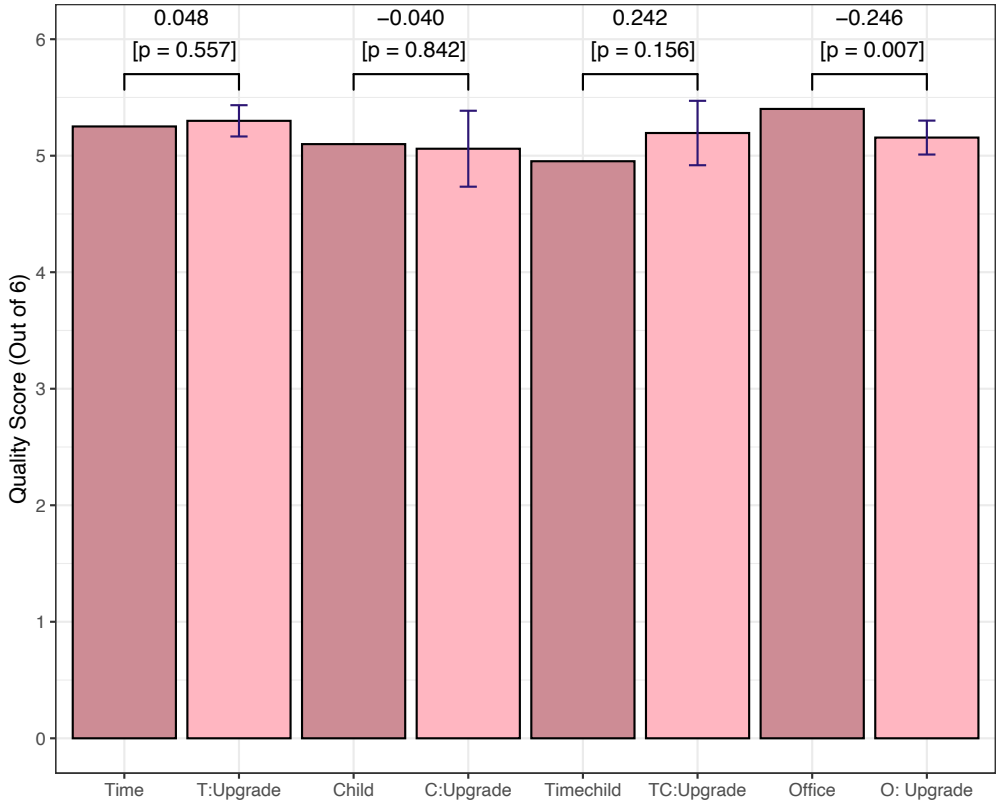
Notes: This figure shows the fraction of workers who stay from weeks 1-4 of the job. Week of job is on the *x*-axis, and retention is on the *y*-axis.

Figure A.18: Average Task Quality

(a) Selection



(b) Treatment Effects

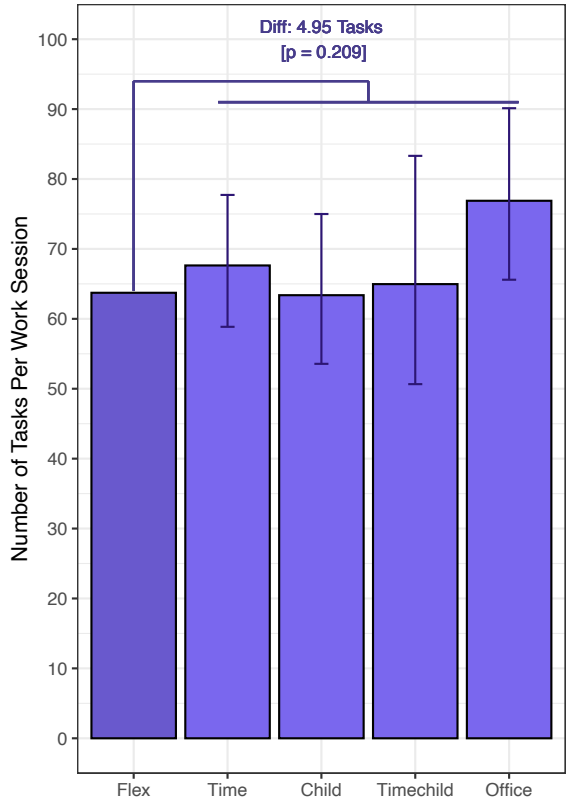


Notes:

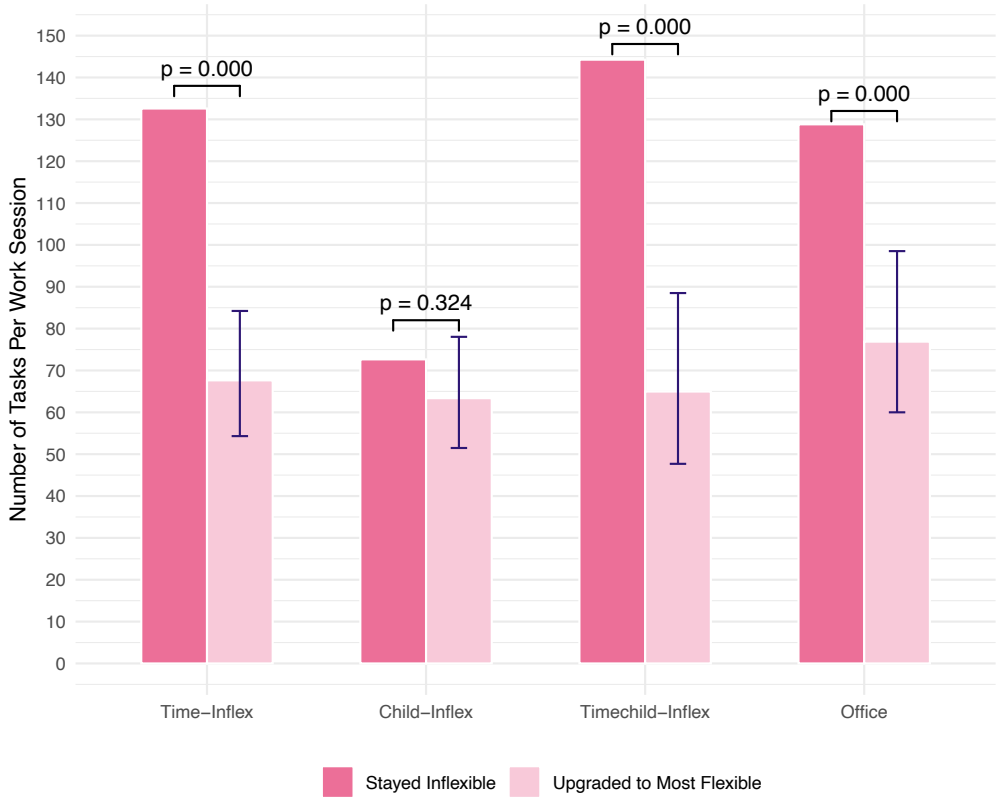
–

Figure A.19: Tasks Per Worksession

(a) Selection



(b) Treatment Effects

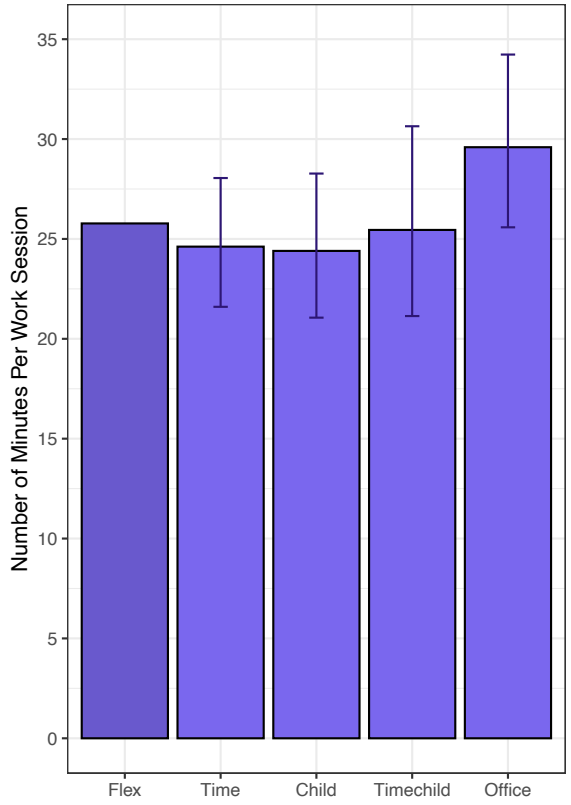


Notes:

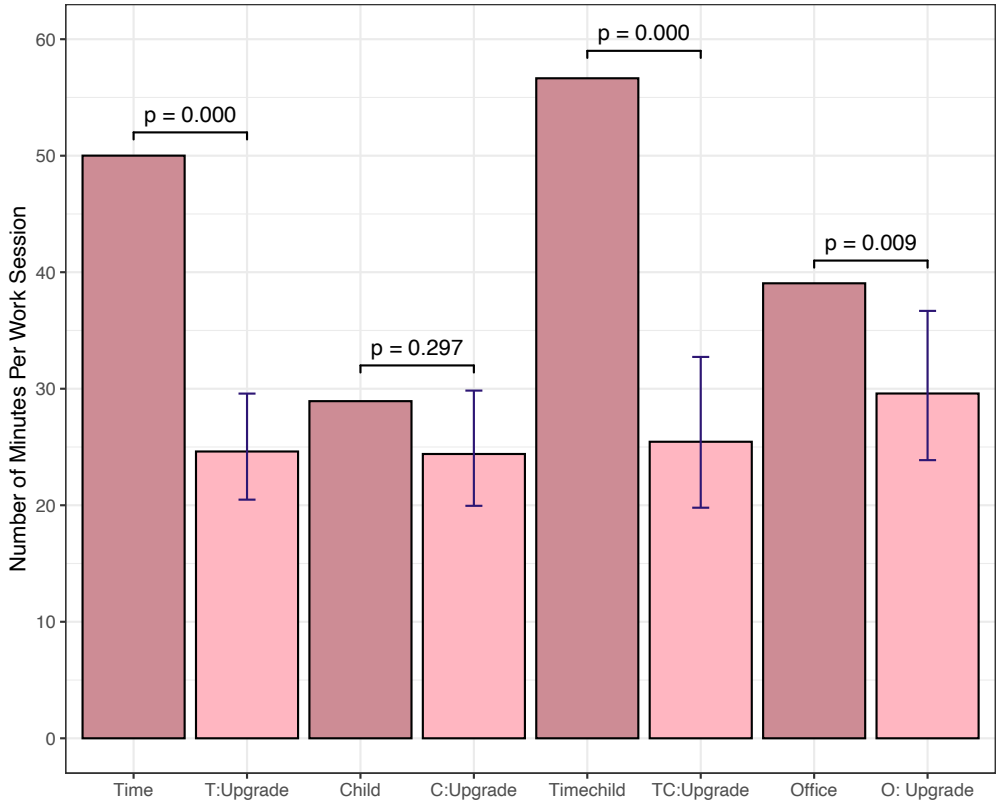
–

Figure A.20: Minutes Per Worksession

(a) Selection



(b) Treatment Effects



Notes:

–

Figure A.21

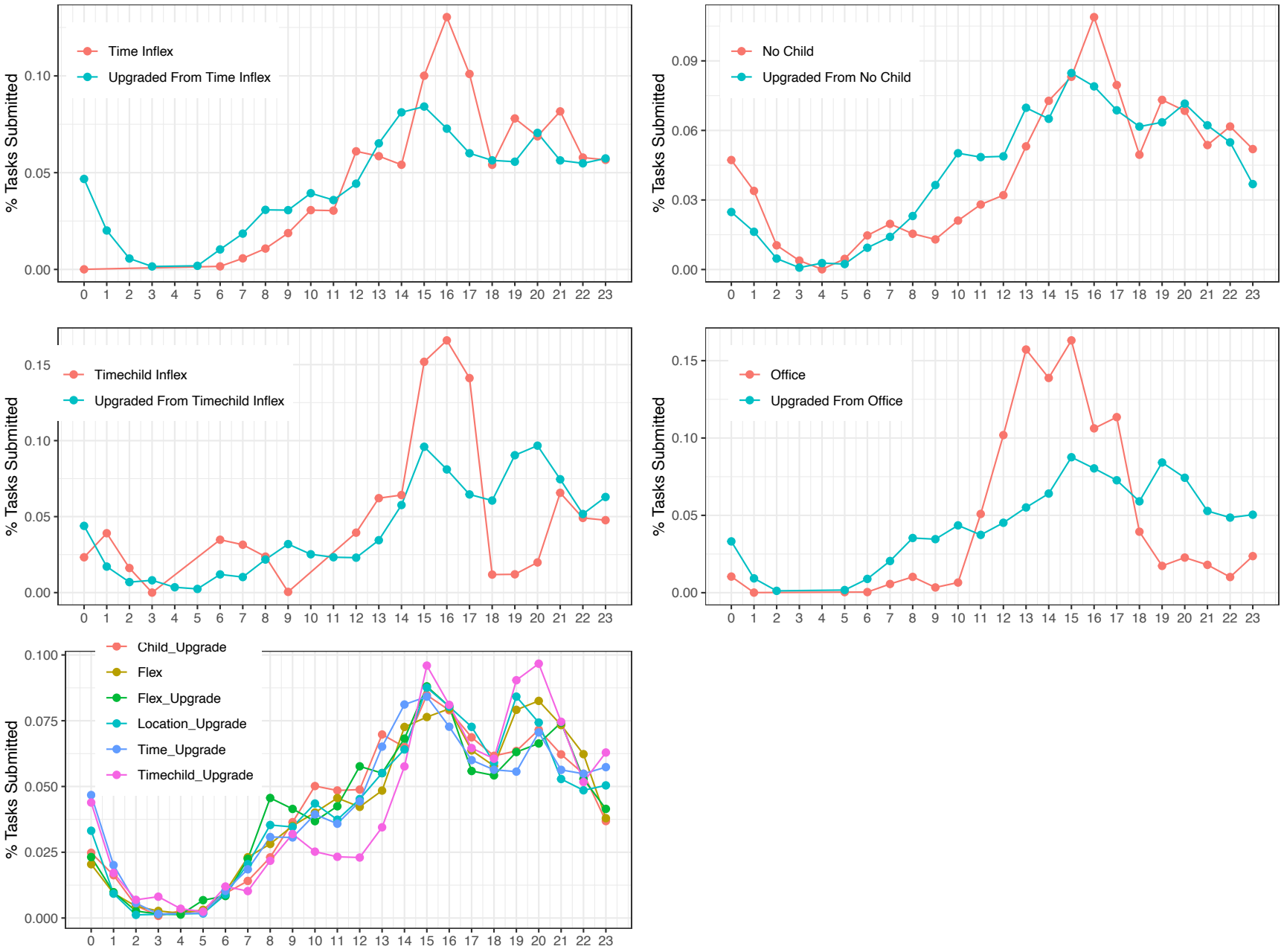
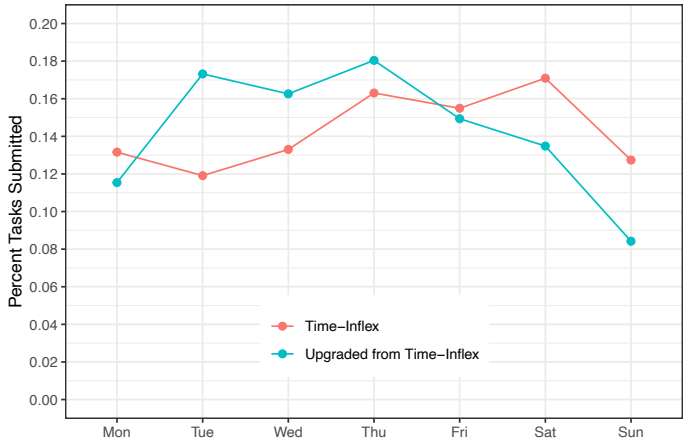
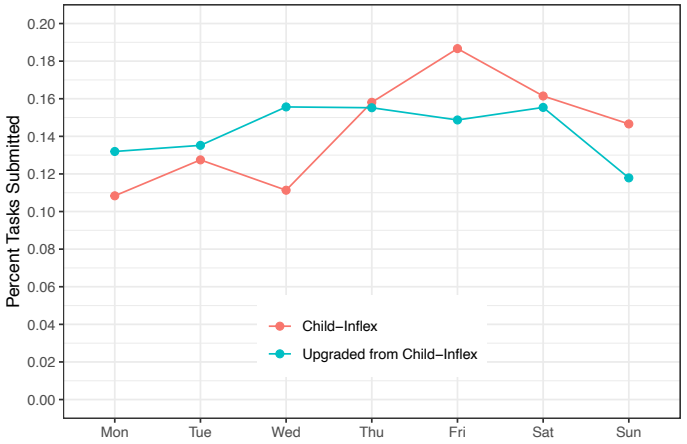
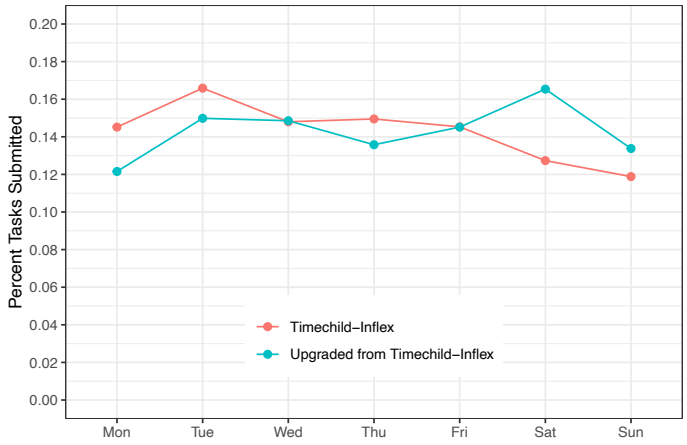
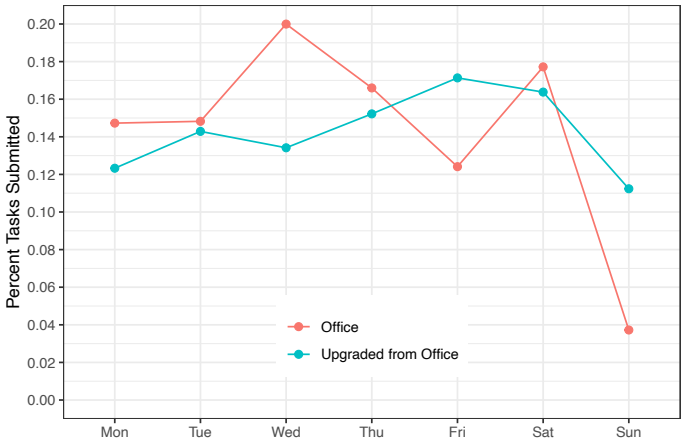

Time Inflex
Upgraded From Time Inflex
% Tasks Submitted
No Child
Upgraded From No Child
Timechild Inflex
Upgraded From Timechild Inflex
Office
Upgraded From Office
Child_Upgrade
Flex
Flex_Upgrade
Location_Upgrade
Time_Upgrade
Timechild_Upgrade

Figure A.22


Percent Tasks Submitted
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Time−Inflex
Upgraded from Time−Inflex


Percent Tasks Submitted
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Child−Inflex
Upgraded from Child−Inflex


Percent Tasks Submitted
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Timechild−Inflex
Upgraded from Timechild−Inflex


Percent Tasks Submitted
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Office
Upgraded from Office

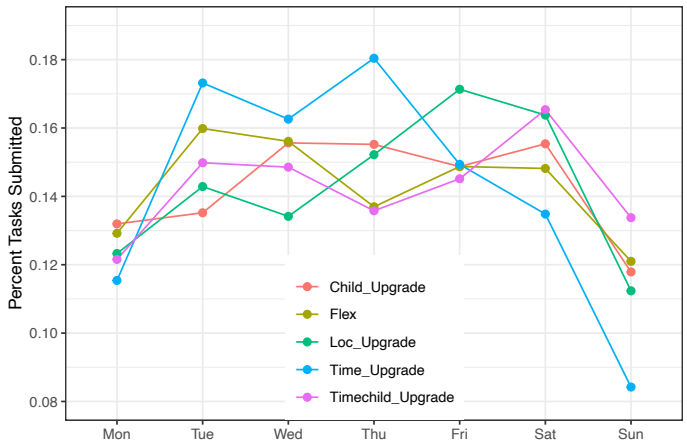

Percent Tasks Submitted
0.18
0.16
0.14
0.12
0.10
0.08
Mon
Tue
Wed
Thu
Fri
Sat
Sun
Child_Upgrade
Flex
Loc_Upgrade
Time_Upgrade
Timechild_Upgrade

Figure A.23

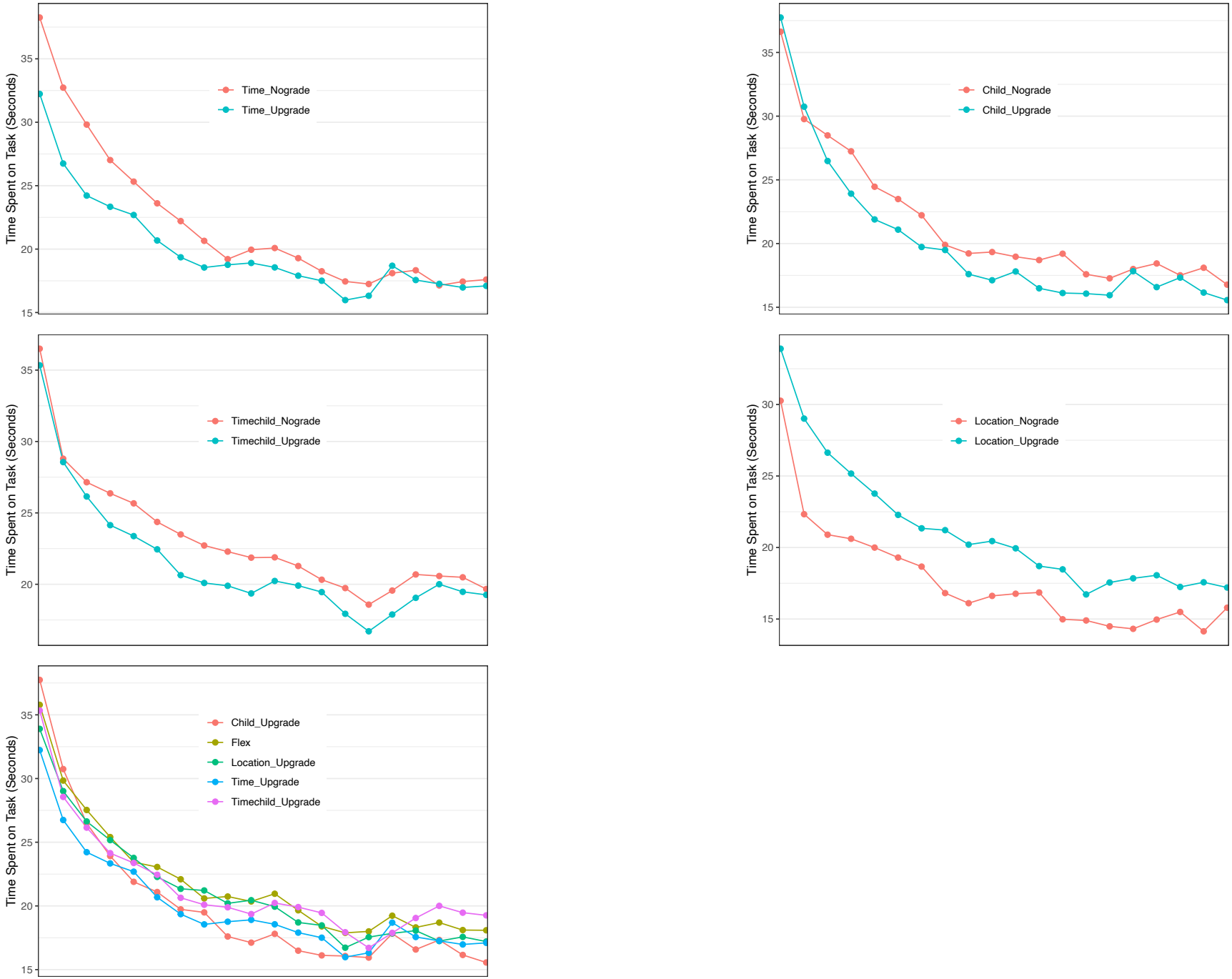

Time Spent on Task (Seconds)
Time_Nograde
Time_Upgrade
Child_Nograde
Child_Upgrade
Timechild_Nograde
Timechild_Upgrade
Location_Nograde
Location_Upgrade
Child_Upgrade
Flex
Location_Upgrade
Time_Upgrade
Timechild_Upgrade

Figure A.24: Monotonicity of job take up in flexibility

Did not accept job below, but accepted…	1 less flexible job	2 less flexible jobs	3 less flexible jobs
Most Flexible	<1%	<1%	<1%
Time-Inflexible	4.4%	1.7%	<1%
Child-Inflexible	4.8%	<1%	-
Timechild-Inflexible	6.0%	-	-

Figure A.25: Effect on perceptions of talent and abilities are driven by women who worked

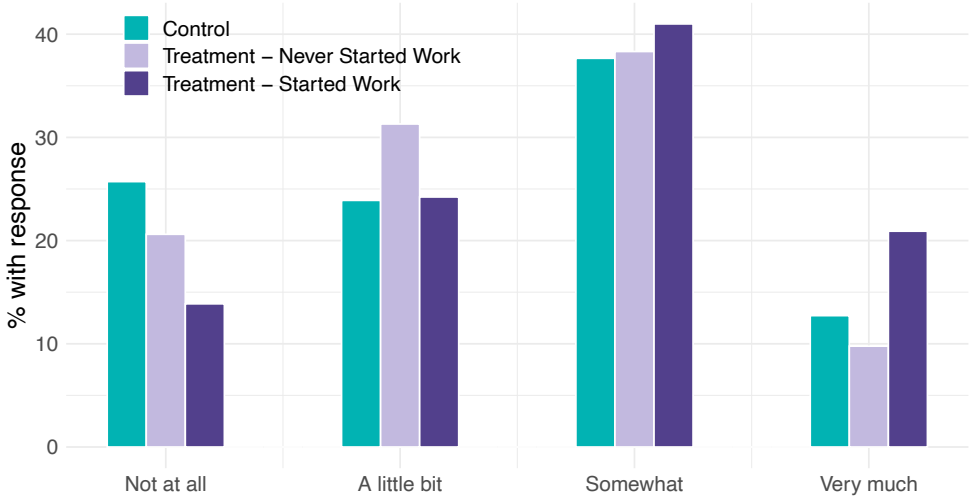


Figure A.26: “Flow” effects over the course of the worksession

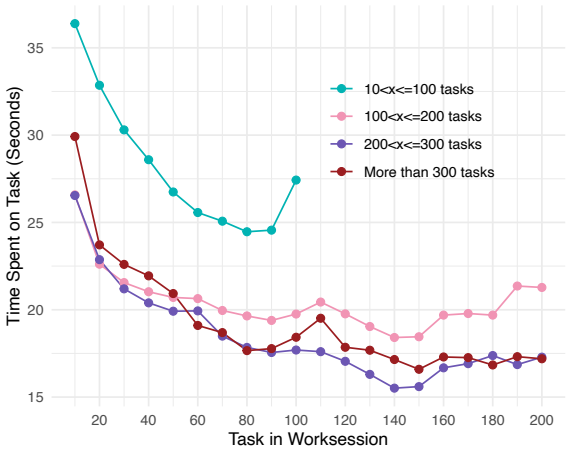
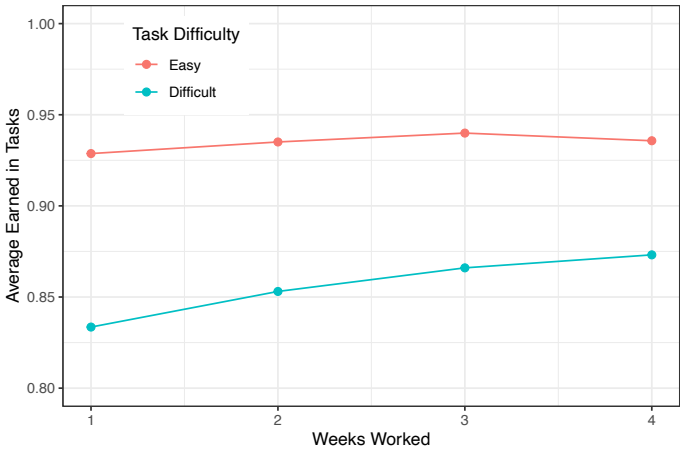
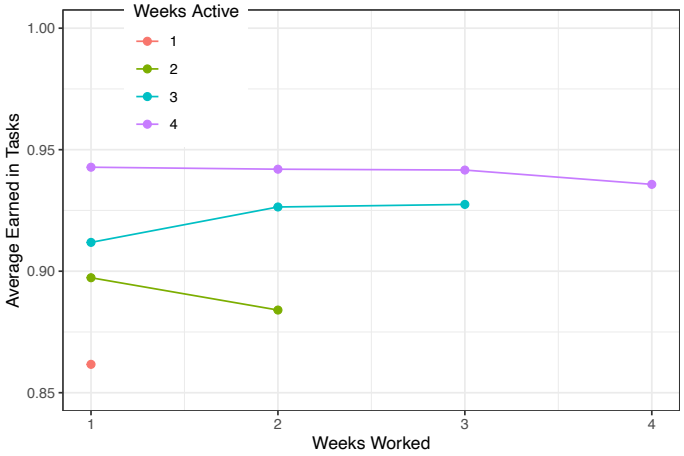


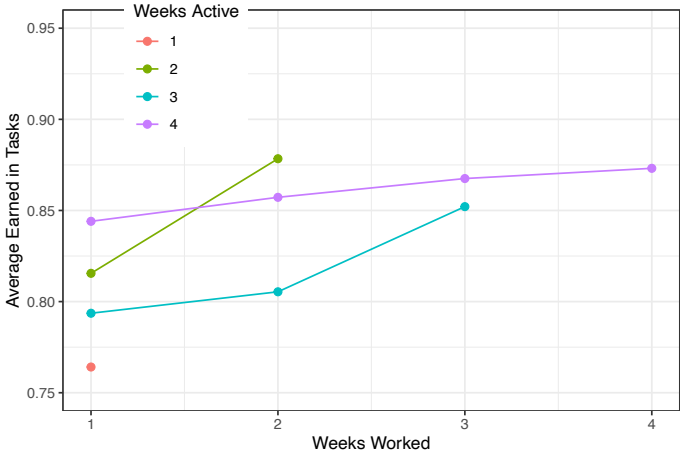
Figure A.27: Learning over weeks



Easy tasks:



Difficult tasks:



Appendix Tables

Table A.8: Balance Table for Participants Who Were Randomized

	(1) All Participants	(2) Control	(3) Treatment	(2)-(3) Pairwise t-test
Age	29.911	29.760	29.962	0.592
Completed 10th Standard (=1)	0.493	0.510	0.488	0.617
Scheduled Caste/Tribe (=1)	0.385	0.376	0.388	0.968
Hindu (=1)	0.754	0.724	0.764	0.094*
Never Married (=1)	0.069	0.062	0.071	0.298
Number HH Members	4.626	4.702	4.601	0.332
Parent-in-Law in HH (=1)	0.402	0.433	0.392	0.119
Has Child Under 8 (=1)	0.478	0.471	0.481	0.185
Job Decision Final Say (=Self)	0.359	0.376	0.353	0.382
Has Own Smartphone (=1)	0.725	0.731	0.723	0.366
Gender Attitudes Index	0.025	-0.000	0.034	0.112
Agency Index	0.013	-0.000	0.018	0.674
Number of participants	1670	420	1250	
F-test of joint significance (F-stat)				1.152
F-test, number of observations				1670

Notes: The data in this table are from women's baseline surveys and compares participants who were randomized into the control group versus one of the treatment groups. All job treatment groups are pooled in this table. The regressions include strata fixed effects.

Table A.9: Effect of job attributes on take up (with controls)

Dependent Variables:	Baseline	Job Offer	Start Work
Model:	(1)	(2)	(3)
Variables			
Time inflexible	-0.04***	-0.04	-0.01
	(0.01)	(0.04)	(0.05)
Child inflexible	-0.27***	-0.17***	-0.20***
	(0.01)	(0.04)	(0.05)
Time & child inflexible	-0.32***	-0.13***	-0.17***
	(0.01)	(0.04)	(0.05)
Location inflexible	-0.42***	-0.25***	-0.34***
	(0.01)	(0.04)	(0.03)
Fixed-effects			
strata_control	Yes	Yes	Yes
Fit statistics			
Observations	8,290	1,250	1,250
Flex Take Up	0.98	0.75	0.48

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work (i.e. submitted job tasks to the employer for review). At baseline, each participant was asked about each of the five work arrangements (in randomized order). At the job offer and starting work stage, each participant had been randomized to one work arrangement. The regressors are four dummy variables, one for each inflexible work arrangement. The omitted group represents the percentage of participants who took up the most flexible job. These regressions control for strata fixed effects and the following participant covariates: age, marital status, employment history, educational attainment, cohabitation with parents-in-law, religion, and household members.

Table A.10: Heterogeneous Impact of Flexibility on Take-Up by Who Has Final Say in Labor Supply

	Baseline (1)	Job Offer (2)	Start Work (3)
Flexible Job	0.36***	0.18***	0.29***
	(0.01)	(0.04)	(0.03)
Final Say in Job Decision = Self	0.14***	0.09**	0.15***
	(0.02)	(0.04)	(0.04)
Flexible Job × Final Say in Job Decision = Self	-0.12***	0.01	-0.08
	(0.02)	(0.06)	(0.06)
Inflexible Take-Up, Final Say Not Self	0.59	0.51	0.16
P-val: equality of coefficients			
Flexible Job: Self Has Final Say = Other Has Final Say	0.013	0.026	0.051
Observations	8,290	1,250	1,250

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The third row of coefficients is the (negative) differential importance of flexibility for participants who say on the baseline survey that they have the final say in their own labor supply.

Table A.11: Heterogeneous Impact of Flexibility on Take-Up by Having Young Child

	Baseline (1)	Job Offer (2)	Start Work (3)
Flexible Job	0.22***	0.16***	0.22***
	(0.01)	(0.04)	(0.04)
Has Child Under 8	-0.20***	-0.06*	-0.09**
	(0.02)	(0.03)	(0.04)
Flexible Job × Has Child Under 8	0.19***	0.05	0.08
	(0.02)	(0.06)	(0.05)
Inflexible Take-Up, No Child Under 8	0.74	0.57	0.26
P-val: equality of coefficients			
Flexible Job: Has Child = No Child	0.394	0.793	0.770
Observations	8,290	1,250	1,250

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The third row of coefficients is the differential importance of flexibility for participants who have a child under eight years old.

Table A.12: Heterogeneous Impact of Flexibility on Take-Up by Educational Attainment

	Baseline (1)	Job Offer (2)	Start Work (3)
Flexible Job	0.33***	0.19***	0.22***
	(0.01)	(0.04)	(0.03)
>10th Standard	0.03*	0.11***	0.12***
	(0.02)	(0.03)	(0.04)
Flexible Job × >10th Standard	-0.03*	-0.02	0.07
	(0.02)	(0.06)	(0.05)
Inflexible Take-Up, <10th Standard	0.70	0.52	0.20
P-val: equality of coefficients			
Flexible Job: <10th == >10th	0.782	0.023	0.000
Observations	8,290	1,250	1,250

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The third row of coefficients is the additional importance of flexibility for participants who finished at least 10th standard.

Table A.13: Heterogeneous Impact of Flexibility on Take-Up by Household Income

	Baseline (1)	Job Offer (2)	Start Work (3)
Flexible Job	0.33***	0.20***	0.24***
	(0.01)	(0.04)	(0.04)
High Income	0.05***	0.05	0.03
	(0.02)	(0.04)	(0.04)
Flexible Job × High Income	-0.03	-0.02	0.04
	(0.02)	(0.06)	(0.05)
Inflexible Take-Up, Low Income	0.62	0.52	0.21
P-val: equality of coefficients			
Flexible Job: High Income == Low Income	0.006	0.585	0.067
Observations	7,860	1,185	1,185

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The third row of coefficients is the differential importance of flexibility for participants who with household incomes higher than the median in the study.

Table A.14: Heterogeneous Impact of Flexibility on Take-Up by Cohabitation with Parent(s)-in-Law

	Baseline (1)	Job Offer (2)	Start Work (3)
Flexible Job	0.31***	0.22***	0.25***
	(0.01)	(0.03)	(0.03)
Lives with In-Laws	-0.03	0.02	-0.03
	(0.02)	(0.04)	(0.04)
Flexible Job × Lives with In-Laws	0.02	-0.11*	0.02
	(0.02)	(0.06)	(0.05)
Inflexible Take-Up, Not Living with In-Laws	0.65	0.54	0.23
P-val: equality of coefficients			
Flexible Job: With In-Laws == Without In-Laws	0.354	0.042	0.702
Observations	8,290	1,250	1,250

Note: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The third row of coefficients is the additional importance of flexibility for participants who live with at least one parent-in-law.

Table A.15: Heterogeneous Impact of Flexibility on Take-Up by Baseline Gender Attitudes (Continuous)

	Baseline	Job Offer	Start Work
	(1)	(2)	(3)
Flexible Job	0.32***	0.18***	0.25***
	(0.01)	(0.03)	(0.03)
Baseline Gender Attitudes (Continuous)	0.11***	0.19***	0.22***
	(0.02)	(0.04)	(0.05)
Flexible Job × Baseline Gender Attitudes (Continuous)	-0.08***	-0.04	-0.11*
	(0.02)	(0.07)	(0.06)
Standard-Errors	hid	Heteroskedasticity-robust	
Observations	8,290	1,250	1,250
Inflexible Take-Up	0.70	0.52	0.20

Notes: This table presents the results of regressions where the outcome variable is a dummy variable equal to 1 if the participant took up the job offer. Each column shows a different way of measuring job takeup. Column (1) measures take up according to whether or not the participant said they would accept the job if offered it during the baseline survey. Column (2) measures take up according to whether the participant said yes when actually called with a job offer by the jobs team. Column (3) measures take up according to whether or not the participant actually began work. The second row of coefficients is the increased job take up rate for participants who are 1 SD less traditional than average. The third row of coefficients is the (negative) additional importance of flexibility for participants who are 1 SD less traditional.

Table A.16: Effect of Job Treatments (Pooled) on Gender-Related Attitudes

Dependent Variables:	All Gender Attitudes	HH Roles	Women & Work	Technology
Model:	(1)	(2)	(3)	(4)
Variables				
Treatment	0.10*	0.05	0.08	0.09*
	(0.05)	(0.05)	(0.05)	(0.06)
Baseline Gender Attitudes	0.91***	0.79***	0.48***	0.66***
	(0.06)	(0.05)	(0.06)	(0.06)
Fixed-effects				
Strata FE	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,522	1,522	1,522	1,522
Control Mean	0.00	0.00	0.00	0.00

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: Column (1) is an index which includes all 15 gender-related attitudes questions, computed as sum of responses divided by number of questions answered by the respondent. The responses for each question range from 0-3 (strongly disagree, disagree, agree, strongly agree). The regressions include controls for age, whether currently married, whether ever employed before, whether completed 10th standard, whether living with in-laws, whether Hindu, number of household members, and baseline gender-related attitudes. Regressions also includes strata fixed effects, where strata are determined by region, whether participants own their own smartphone, and whether the participant has a child under the age of eight.

Table A.17: Effect of Job Treatments (By Work Arrangement) on Gender Attitudes

	All Gender Attitudes							
	ITT Estimates				2SLS Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
arrangement = Flex	0.08***	0.08***	0.06**	0.06**				
	(0.03)	(0.03)	(0.03)	(0.03)				
arrangement = Time-Inflex	0.11**	0.11**	0.10**	0.10**				
	(0.05)	(0.05)	(0.04)	(0.04)				
arrangement = Child-Inflex	0.08	0.08*	0.06	0.06				
	(0.05)	(0.05)	(0.05)	(0.05)				
arrangement = Time- & Child-Inflex	-0.07	-0.08	-0.06	-0.06				
	(0.06)	(0.06)	(0.05)	(0.05)				
arrangement = Office	0.05	0.05	0.03	0.03				
	(0.04)	(0.04)	(0.04)	(0.04)				
Started Flex					0.15***	0.15***	0.12**	0.12**
					(0.05)	(0.05)	(0.05)	(0.05)
Started Time					0.23**	0.23**	0.22***	0.21**
					(0.10)	(0.10)	(0.08)	(0.09)
Started Child					0.26*	0.26*	0.19	0.18
					(0.15)	(0.15)	(0.14)	(0.14)
Started TimeChild					-0.23	-0.25	-0.15	-0.18
					(0.19)	(0.19)	(0.17)	(0.17)
Started Office					0.34	0.30	0.21	0.18
					(0.25)	(0.25)	(0.23)	(0.23)
Strata Fixed Effects		×		×		×		×
Lasso Selected Controls			×	×			×	×
Observations	1,525	1,525	1,525	1,525	1,525	1,525	1,525	1,525
Control Mean	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table A.18: Effect of Job Treatments (Pooled) on Gender Attitudes: Heterogeneity by Baseline Gender Attitudes

Dependent Variables:	All Gender Attitudes Index	HH Roles	Women & Work	Tech	Ability
Model:	(1)	(2)	(3)	(4)	(5)
Treatment	0.70*	0.19	0.27*	0.18	0.06
	(0.37)	(0.16)	(0.14)	(0.16)	(0.09)
Less Traditional × Treatment	-0.02	0.03	-0.13	0.07	0.02
	(0.43)	(0.19)	(0.15)	(0.18)	(0.10)
Baseline Gender Attitudes	0.44***	0.18***	0.09***	0.11***	0.05***
	(0.04)	(0.01)	(0.01)	(0.02)	(0.01)
Strata FE	Yes	Yes	Yes	Yes	Yes
Observations	1,524	1,524	1,524	1,524	1,524
Control Mean	6.43	2.55	3.87	0.00	0.00

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: Same as previous table, looking at heterogeneous effects by baseline gender attitudes index.

Table A.19: Treatment Effect of Jobs Intervention on Gender Attitudes - Heterogeneity by Baseline Attitudes (Continuous Version)

	Endline Gender Attitudes							
	ITT Estimates				2SLS Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.05**	0.05**	0.05**	0.05**				
	(0.02)	(0.02)	(0.02)	(0.02)				
Treatment × Baseline Gender Attitudes	-0.08	-0.09	-0.10	-0.12*				
	(0.07)	(0.07)	(0.06)	(0.07)				
Started Work					0.14**	0.14**	0.15**	0.15**
					(0.06)	(0.06)	(0.06)	(0.06)
Started Work × Baseline Gender Attitudes					-0.25	-0.28	-0.33*	-0.36**
					(0.18)	(0.18)	(0.18)	(0.18)
Baseline Gender Attitudes	0.45***	0.43***	0.26***	0.26***	0.45***	0.43***	0.26***	0.26***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Strata Fixed Effects		×		×		×		×
Lasso Selected Controls			×	×			×	×
Observations	1,525	1,525	1,525	1,525	1,525	1,525	1,525	1,525
Control Mean (Endline)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Notes: This table tests for the impact of the job treatments on gender attitudes, examining heterogeneity by baseline attitudes. In this analysis, all treatment groups (work arrangements) are pooled together. Gender attitudes are computed as a standardized, weighted average of questions from the baseline survey or the endline survey, in which the weights take into account the covariance structure of the components (Anderson, 2008). Both baseline and endline gender attitudes are included in this analysis as this continuous index variable with mean zero and standard deviation one. The dependent variable is gender attitudes measured on the endline survey, which took place 1-2 weeks after the end of the job treatments. Columns (1)-(4) present intent-to-treat (ITT) estimates, while columns (5)-(8) present two-stage least squares (2SLS) estimates. Heteroskedasticity- robust standard errors are reported in parentheses. Stars next to coefficients represent unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table A.20: Effect of Job Treatments on Women's Agency

	Agency Index		
	(1)	(2)	(3)
Panel A: Treatments Pooled			
Treatment	0.04	0.04	0.02
	(0.04)	(0.04)	(0.03)
Control Mean	-0.01	-0.01	-0.01
Panel B: Heterogeneity by Work Arrangement			
Arrangement = Flex or Time-Inflex	0.05	0.05	0.03
	(0.04)	(0.04)	(0.04)
Arrangement = Child- or Child- & Time-Inflex	0.03	0.03	0.01
	(0.06)	(0.06)	(0.05)
Arrangement = Office	0.00	0.00	-0.03
	(0.06)	(0.06)	(0.05)
Control Mean	-0.01	-0.01	-0.01
Panel C: Heterogeneity by Baseline Agency			
Treatment	0.03	0.03	0.03
	(0.05)	(0.05)	(0.04)
Treatment × Baseline Agency	-0.01	-0.01	-0.02
	(0.07)	(0.07)	(0.05)
Control Mean, Low Baseline Agency	-0.18	-0.18	-0.18
Observations	1,525	1,525	1,525
Strata FE fixed effects		✓	✓
Lasso Selected Controls			✓

Notes: This table presents results about the treatment effect of the jobs intervention on participants' agency. In this analysis, all treatment groups are pooled and compared to the control group. The dependent variable in column (1) is a weighted average of seven questions on the endline survey, while the dependent variables in columns (2)-(5) are subsets of the variables in the column (1) index. The weights on the different index components is informed by their covariance, as in Anderson (2008). Column (2) is a binary variable equal to one if the participant names herself as the person who would have the final say in her own labor supply. Column (3) is a physical mobility index composed of three questions (how often the participant leave homes alone, the participant's ability to leave home without asking permission, and the participant's ability to meet friends without permission). Column (4) is an individual purchases index composed of two questions (to what extent the participant can purchase clothes independently, and to what extend she can buy things from the market without asking). Column (5) is the standardized variable for how much of a say the participant has in significant household purchases. The sample size for the sub-indices varies depending on whether participants indicated that the question was relevant to them (e.g. for column (4) whether or not they had made any purchases in the last month, and for column (5) whether their household had made any significant purchases in the last month). All regressions include lasso-selected controls and strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. Stars next to coefficients represent unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table A.21: Effect of Job Treatments on Women's Psychological Well-Being

	Psych Index (1)	Sleep Peacefully (2)	(Not) Overwhelmed (3)	Happy (4)	(Not) Worried (5)
Panel A: Treatments Pooled					
Treatment	0.02	-0.02	0.07	-0.02	-0.01
	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Control Mean	0.00	0.00	0.00	0.00	0.00
Panel B: Heterogeneity by Work Arrangement					
Arrangement = Flex or Time	0.03	0.03	0.10	-0.02	-0.01
	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Arrangement = Child or Timechild	-0.01	-0.12	0.01	0.04	-0.07
	(0.05)	(0.09)	(0.09)	(0.08)	(0.09)
Arrangement = Office	-0.02	-0.13	0.02	-0.04	0.03
	(0.06)	(0.09)	(0.09)	(0.09)	(0.09)
Panel C: Heterogeneity by Baseline Wellness					
Treatment	0.02	-0.06	0.09	0.03	-0.02
	(0.05)	(0.07)	(0.07)	(0.07)	(0.08)
Treatment $\times$ >Well at Baseline	0.00	0.08	-0.04	-0.08	0.02
	(0.06)	(0.09)	(0.08)	(0.08)	(0.09)
Control Mean, < Well at Baseline	-0.16	-0.15	-0.10	-0.21	-0.21
Observations	1,524	1,524	1,524	1,524	1,524
Strata FE	✓	✓	✓	✓	✓
Lasso Selected Controls	✓	✓	✓	✓	✓

Notes: This table presents results about the treatment effect of the jobs intervention on participants' psychological wellbeing. In this analysis, all treatment groups are pooled and compared to the control group. The dependent variable in column (1) is a weighted average of four questions on the endline survey, while the dependent variables in columns (2)-(5) are the components making up the column (1) index. The weights on the different index components in column (1) are informed by their covariance, as in Anderson (2008). For the questions corresponding to columns (2)-(5), participants were asked about how often they felt a certain way in the last month (i.e. during the treatment). Column (2) is how often they slept peacefully, column (3) is how often they were generally feeling happy, column (4) is how often they were feeling anxious, and column (5) is how often they felt overwhelmed. The outcomes in columns (4) and (5) are negated so that positive values always correspond to better psychological wellbeing. In response to these questions, participants could answer that in the last month they felt this way (i) Never, (ii) A few days, (iii) Around half the days, (iv) More than half the days, and (v) Nearly every day. The answers are all standardized to have mean zero and standard deviation equal to one. All regressions include lasso-selected controls and strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. Stars next to coefficients represent unadjusted p-values (* significant at 10%; ** at 5%; *** at 1%).

Table A.22: Treatment Effect on Take Up of Digital vs Non-Digital Jobs

	Started Work in Round 2					
	R2: Flexible		R2: Inflexible		Interaction	
	Digital	Non-Digital	Digital	Non-Digital	Both	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.09	-0.09	0.01	0.06	-0.09	0.05
	(0.08)	(0.09)	(0.06)	(0.08)	(0.08)	(0.07)
R2 = Digital					0.02	-0.11
					(0.10)	(0.08)
Treatment × R2 = Digital					0.03	-0.05
					(0.11)	(0.09)
Observations	218	218	217	218	436	435
Strata Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Lasso Selected Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes:

Table A.23: Spillovers to Control Group Job Take Up

	Round 2: Started Work					
	R2: Any Job			R2: Flexible Jobs		
	(1)	(2)	(3)	(4)	(5)	(6)
Num Friends (Flexible Jobs)	0.05			0.16**		
	(0.05)			(0.07)		
	[0.333]			[0.018]		
>= 1 Friend (Flexible Jobs)		0.04			0.15	
		(0.08)			(0.15)	
		[0.598]			[0.305]	
Fraction Friends (Flexible Jobs)			0.03			0.10
			(0.08)			(0.13)
			[0.693]			[0.439]
Num Friends (Any Group)	-0.01	0.01	0.02	-0.07	-0.01	0.01
	(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.04)
	[0.879]	[0.703]	[0.608]	[0.167]	[0.893]	[0.841]
Strata Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Lasso Selected Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	367	367	367	149	149	149
Control, No Friends Take Up	0.43	0.43	0.43	0.64	0.64	0.64

Notes: This table presents results on spillovers to the control group in terms of starting work during the second round of jobs.

- These regressions only include participants who were in the control group.
- The dependent variable is a binary variable equal to one if the participant started the job that she was assigned to during Jobs Round 2. This outcome is regressed on the participant's total number of friends in the study, along with a measure of these friends' exposure to the flexible jobs (most flexible and time-inflexible, the arrangements with the highest job take up).
- Exposure is measured in three ways: (1) number of friends assigned to flexible jobs, (2) a binary variable equal to one if at least one friend was was assigned to a flexible job, and (3) the fraction of a participant's friends who were assigned to a flexible job.
- Friend lists are collected at baseline as well as during a follow-up survey 4-6 months after the endline survey. In this table, we include any friend who was listed on the baseline survey, as well as friends listed on the follow-up survey who (i) has been friends with the participant for more than one year, and (ii) is a relative or neighbor of the participant, to alleviate concerns that they met these friends through the study.
- Columns (1)-(3) consider the take up of any job arrangements in Round 2, while columns (4)-(6) consider the take up of flexible jobs in Round 2.
- Standard errors in parentheses $(\cdot)$ are robust to heteroskedasticity. Brackets $[\cdot]$ report unadjusted p-values, as do stars next to coefficients (* significant at 10%; ** at 5%; *** at 1%).

Table A.24: Spillover Effects (Flexible & Inflexible Friends) on Round 2 Job Take Up

Started Work in Round 2						
	All Arrangements			Flex or Time		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Friends (Flexible Jobs)	0.09			0.20***		
	(0.06)			(0.07)		
Number Friends (Inflexible Jobs)	0.07*			0.07		
	(0.04)			(0.04)		
>=1 Friend (Flexible Jobs)		0.04			0.15	
		(0.09)			(0.15)	
>=1 Friend (Inflexible Jobs)		0.00			0.01	
		(0.09)			(0.14)	
Fraction Friends (Flexible Jobs)			0.03			0.10
			(0.08)			(0.13)
Fraction Friends (Inflexible Jobs)			0.02			0.01
			(0.06)			(0.08)
Number Friends (Any Group)	-0.05	0.01	0.02	-0.11**	-0.01	0.01
	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)
Strata Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Lasso Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	367	367	367	149	149	149
Control, No Friends	0.43	0.43	0.43	0.39	0.39	0.39

Notes:

Table A.25: Effect of Job Treatments (Pooled) on Children's Household Help

	Help Cook	Help Clean	Help Childcare	Dad Helps (=1) (More Than Never)
	(1)	(2)	(3)	(4)
Treatment	0.01	0.10	0.03	0.09**
	(0.11)	(0.11)	(0.15)	(0.05)
	[0.924]	[0.355]	[0.853]	[0.045]
Control Mean (Endline)	1.13	1.09	2.47	0.58
Heterogeneity by Child Age				
Treatment	0.15	0.26*	-0.13	0.17***
	(0.14)	(0.14)	(0.19)	(0.06)
	[0.284]	[0.064]	[0.494]	[0.005]
Treatment $\times < 12$	-0.29	-0.31	0.36	-0.19**
	(0.23)	(0.23)	(0.30)	(0.09)
	[0.201]	[0.168]	[0.237]	[0.037]
< 12	-0.04	-0.17	-0.57**	0.19**
	(0.20)	(0.20)	(0.25)	(0.08)
	[0.826]	[0.388]	[0.026]	[0.016]
Control Mean, > 12	1.39	1.38	2.51	0.51
Heterogeneity by Child Gender				
Treatment	-0.22	-0.02	-0.01	0.15**
	(0.17)	(0.17)	(0.19)	(0.07)
	[0.211]	[0.925]	[0.958]	[0.029]
Treatment $\times$ Male	0.40*	0.17	0.06	-0.10
	(0.22)	(0.22)	(0.28)	(0.20)
	[0.076]	[0.439]	[0.820]	[0.274]
Male	-0.48**	-0.49**	-0.15	0.13
	(0.20)	(0.20)	(0.24)	(0.08)
	[0.016]	[0.013]	[0.536]	[0.104]
Control Mean, Female	1.39	1.38	2.51	0.51
Observations	606	606	315	603
Strata Fixed Effects	$\times$	$\times$	$\times$	$\times$
Baseline Behavior	$\times$	$\times$	$\times$	$\times$

Table A.26: Effect of Job Treatments (Pooled) on Children's Aspirations

	Aspire - UGrad (1)	Aspire - Masters (2)
Treatment	0.00	-0.03
	(0.03)	(0.04)
Control Mean (Endline)	0.88	0.34
Heterogeneity by Child Age		
Treatment	-0.04	0.01
	(0.04)	(0.06)
Treatment $\times$ Age $<$ 12	0.09	-0.08
	(0.06)	(0.09)
Age $<$ 12	-0.06	0.04
	(0.05)	(0.08)
Control Mean, Age $>$ 12	0.91	0.33
Heterogeneity by Child Gender		
Treatment	0.00	0.01
	(0.04)	(0.07)
Treatment $\times$ Male	-0.02	-0.07
	(0.06)	(0.09)
Male	-0.01	-0.02
	(0.05)	(0.08)
Control Mean, Female	0.89	0.36
Observations	602	602

Note: The outcome variable in column (1) is a dummy variable equal to 1 if the child says they would like to finish at least an undergraduate degree. The outcome variable in column (1) is a dummy variable equal to 1 if the child says they would like to finish at least a master's degree.

Table A.27: Quality of Tasks (Selection)

	Quality Index (1)	Accuracy (2)	Volume (3)	Fluency (4)	Earned in Task (5)
Accept Time-Inflex, Then Upgrade	0.03	0.00	0.02	0.02	0.01
	(0.12)	(0.04)	(0.04)	(0.05)	(0.01)
Accept Child-Inflex, Then Upgrade	-0.10	-0.04	-0.01	-0.05	-0.01
	(0.13)	(0.05)	(0.04)	(0.05)	(0.01)
Accept Timechild-Inflex, Then Upgrade	0.00	-0.02	0.02	0.00	0.00
	(0.09)	(0.03)	(0.03)	(0.04)	(0.01)
Accept Office, Then Upgrade	0.02	0.00	0.03	-0.01	0.00
	(0.08)	(0.03)	(0.03)	(0.03)	(0.01)
Observations	949,543	949,543	949,543	949,543	949,543
Flexible Mean	1.79	1.83	1.55	5.17	0.93
Task type fixed effects	✓	✓	✓	✓	✓
Week fixed effects	✓	✓	✓	✓	✓

Notes:

Table A.28: Quality of Tasks (Impact)

	Quality Index (1)	Accuracy (2)	Volume (3)	Fluency (4)	Earned in Task (5)
Time-Inflexible					
Upgraded to Flex	-0.05	-0.02	-0.02	-0.01	0.00
	(0.13)	(0.05)	(0.04)	(0.05)	(0.01)
Child-Inflexible					
Upgraded to Flex	-0.03	0.24***	-0.12***	-0.15*	0.05**
	(0.19)	(0.09)	(0.05)	(0.08)	(0.02)
Timechild-Inflexible					
Upgraded to Flex	0.20	0.39***	-0.12***	-0.07	0.08***
	(0.16)	(0.09)	(0.02)	(0.08)	(0.02)
Office					
Upgraded to Flex	-0.20***	-0.08***	-0.05**	-0.08*	-0.03***
	(0.08)	(0.02)	(0.02)	(0.04)	(0.01)
Observations	273,942	273,942	273,942	273,942	273,942
Flexible Mean	1.79	1.83	1.55	5.17	0.93
Task Type fixed effects	✓	✓	✓	✓	✓
Week fixed effects	✓	✓	✓	✓	✓

Notes:

Table A.29: Quality of Tasks By Task Difficulty (Selection)

	Quality Index (1)	Accuracy (2)	Volume (3)	Fluency (4)	Earned in Task (5)
Accept Time-Inflex, Then Upgrade	0.04	0.00	0.02	0.02	0.01
	(0.12)	(0.04)	(0.04)	(0.05)	(0.01)
Accept Child-Inflex, Then Upgrade	-0.06	-0.02	0.00	-0.04	0.00
	(0.12)	(0.04)	(0.04)	(0.05)	(0.01)
Accept Timechild-Inflex, Then Upgrade	0.04	0.00	0.03	0.01	0.00
	(0.08)	(0.03)	(0.03)	(0.04)	(0.01)
Accept Office, Then Upgrade	0.02	0.01	0.03	-0.01	0.01
	(0.07)	(0.02)	(0.03)	(0.03)	(0.01)
Difficult Task	-0.52***	-0.26***	-0.11***	-0.15***	-0.07***
	(0.07)	(0.04)	(0.02)	(0.02)	(0.01)
Difficult Task × Accept Time-Inflex, Then Upgrade	-0.03	-0.01	0.00	-0.02	0.00
	(0.13)	(0.06)	(0.03)	(0.03)	(0.02)
Difficult Task × Accept Child-Inflex, Then Upgrade	-0.13	-0.07	-0.04	-0.03	-0.02
	(0.15)	(0.07)	(0.04)	(0.04)	(0.02)
Difficult Task × Accept Timechild-Inflex, Then Upgrade	-0.12	-0.06	-0.03	-0.03	-0.02
	(0.12)	(0.06)	(0.03)	(0.03)	(0.02)
Difficult Task × Accept Office, Then Upgrade	0.01	-0.01	0.01	0.01	0.00
	(0.10)	(0.05)	(0.03)	(0.03)	(0.01)
Observations	949,543	949,543	949,543	949,543	949,543
Flexible Mean	5.33	1.87	1.86	1.60	0.95
Week fixed effects	✓	✓	✓	✓	✓

Notes:

Table A.30: Quality of Tasks By Task Difficulty (Impact)

	Quality Index (1)	Accuracy (2)	Volume (3)	Fluency (4)	Earned in Task (5)
Time-Inflexible to Flex					
arrangement = Time	0.02	0.01	0.01	0.00	0.00
	(0.12)	(0.04)	(0.04)	(0.05)	(0.01)
Difficult Task	-0.55***	-0.27***	-0.11***	-0.17***	-0.07***
	(0.10)	(0.05)	(0.03)	(0.03)	(0.01)
Difficult Task × arrangement = Time	0.10	0.04	0.01	0.05	0.01
	(0.15)	(0.08)	(0.04)	(0.04)	(0.02)
Child-Inflexible to Flex					
arrangement = Child	0.04	-0.22***	0.09**	0.17**	-0.04*
	(0.17)	(0.08)	(0.04)	(0.08)	(0.02)
Difficult Task	-0.65***	-0.32***	-0.15***	-0.18***	-0.08***
	(0.13)	(0.06)	(0.03)	(0.04)	(0.02)
Difficult Task × arrangement = Child	0.00	-0.07	0.11***	-0.04	-0.01
	(0.17)	(0.09)	(0.04)	(0.06)	(0.02)
Timechild-Inflexible to Flex					
arrangement = Timechild	-0.19	-0.36***	0.09***	0.09	-0.08***
	(0.15)	(0.09)	(0.02)	(0.07)	(0.02)
Difficult Task	-0.64***	-0.32***	-0.14***	-0.18***	-0.08***
	(0.10)	(0.05)	(0.03)	(0.03)	(0.01)
Difficult Task × arrangement = Timechild	-0.04	-0.10	0.12***	-0.06	-0.02
	(0.13)	(0.09)	(0.03)	(0.05)	(0.02)
Office to Flex					
arrangement = Loc	0.10	0.03	0.02	0.06	0.01
	(0.08)	(0.02)	(0.02)	(0.04)	(0.01)
Difficult Task	-0.51***	-0.27***	-0.10***	-0.14***	-0.07***
	(0.07)	(0.04)	(0.02)	(0.02)	(0.01)
Difficult Task × arrangement = Loc	0.33***	0.19***	0.08***	0.06*	0.05***
	(0.10)	(0.05)	(0.02)	(0.03)	(0.01)
Week fixed effects	✓	✓	✓	✓	✓

Notes:

A.1 Spillovers to the control group

There is suggestive evidence that control group participants who had more friends who worked during Round 1 were more likely to start work in Round 2.

Measuring spillovers. To measure exposure to treatment, we estimate the impact of having friends in the treatment group, conditional on total number of friends in the study. We use three methods to measure having friends in the treatment group: (i) number of friends, (ii) having at least one friend, and (iii) fraction of friends assigned to treatment. Because of low take up of work in inflexible jobs, to measure exposure to friends working rather than just receiving a job offer, in our main specification we include only friends assigned to the flexible jobs (Flex or Time), but we also show results of specifications including friends assigned to inflexible jobs in Table A.24.

To collect information on who the participant knew in the study, we ask at both baseline and endline for participants to list anyone else they knew who took part in our surveys. We then ask for their relationship to this person, how long they have known this person for, and how often they speak with each other. Because recruitment and baseline were done on a rolling basis, the initial lists collected at baseline are very incomplete, and we use the endline lists in our main specifications, eliminating anyone who the participant said they met within the last year (i.e. after the start of the study). In order to mitigate concerns that the endline lists could still include people who the participant met as a result of the study, we use only people who are either relatives or neighbors of the participants, with the idea being that it is very unlikely they met a relative or neighbor as a result of the study.

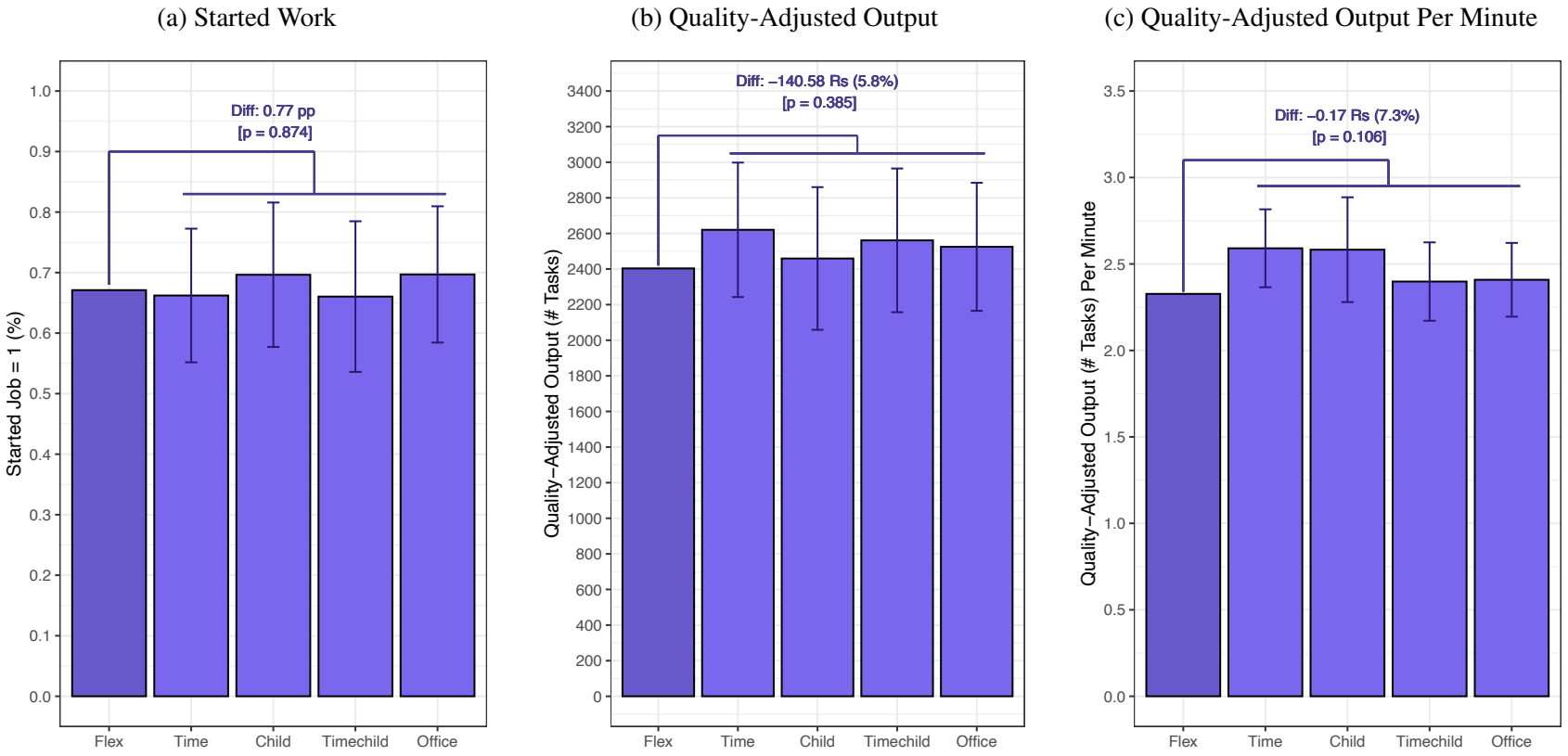
Estimation. We test for the presence of spillovers from the treatment group to the control group using the following regression:

$$\text{TakeUp}_{is} = \gamma_s + \gamma_1 \text{JobExposure}_{is} + \gamma_2 \text{NumFriends}_{is} + \gamma_3 \theta_{is} + \varepsilon_{is} \tag{A.11}$$

where TakeUp_{is} is a dummy variable equal to one if participant i started her job in Round 2; NumFriends_i is the total number of friends that i has in the study; JobExposure_i is one of the three measures of exposure to the treatment group described above; and θ_{is} is a vector of control variables selected using a double LASSO. The coefficient of interest is γ_1, which isolates the causal effect of control participants' exposure to the treatment group on their subsequent job take up. Holding fixed i's total number of friends in the study, i's exposure to the treatment group (e.g. number of friends randomly assigned to the treatment group) is as good as randomly assigned. *Spillover effects.* Overall, the estimates from these regressions suggest that participants with more friends assigned to the flexible jobs

in Round 1 are more likely to take up flexible jobs in Round 2 than participants with less friend exposure. Table A.23 presents results of the spillover regressions, looking at take up of any Round 2 job (columns 1-3) and take up of flexible jobs (columns 4-6). Participants in the control group are 16 pp ($p = 0.018$) more likely to start work in the flexible jobs for each of their friends who was assigned to the flexible jobs, conditional on the number of friends they had in the study. The point estimates are also positive, but smaller and insignificant, for take up of any round 2 job. For the other measures of treatment exposure, the point estimates are similar but less precisely estimated. Results from comparing the effects of friend exposure to the flexible jobs versus inflexible jobs are consistent with the idea that control group participants are more likely to take up jobs if they heard about those jobs from their friends (Table A.24). There are no effects of having friend exposure to the inflexible jobs on take up of flexible jobs in Round 2.

Figure A.28: Effects of Flexibility on Worker Composition



Notes: This figure shows how flexible work arrangements affect worker composition in terms of job performance.

- Three job performance measures are included: In panel A, the likelihood that the worker starts the job after accepting it; in Panel B, the total quality-adjusted output produced by the worker during the month-long job; and in Panel C, a measure of work efficiency that is quality-adjusted output divided by minutes spent actively working. Minutes spent working are measured according to number of unique minutes during which the worker took an action on the job app.
- All workers in these figures worked in the most flexible arrangement. The performance measures are plotted separately for workers initially offered the most flexible job (Flex), and the workers who were initially offered and accepted a less flexible job (Time for time-inflexible, Child for child-inflexible, Timechild for time- and child-inflexible, and Office for office-based) and then were randomly selected for an upgrade to the most flexible job.
- The brackets report the magnitude and p-value for the difference between performance of the Flex group and performance of all the less flexible groups pooled together. Whiskers indicate 90% confidence intervals.

Figure A.29: Effect of flexible work arrangements, including “no shows”