Psychosocial Wellbeing of Rohingya Refugees

Data Collection Report

RTM International

September 27, 2023

Table of Contents

1. Introduction

Recent work has established that employment is a source of great psychosocial value. Yet these returns to employment - both from pecuniary and non-pecuniary mechanisms - are likely to differ by one's role in their social context, a position inseparable from gender identity in much of the world. In Bangladesh, the site of this study, 88% of men and women believe that "when a mother works for pay, the children suffer," and more than 90% believe that men have a greater right to scarce jobs than women.[1] Such norms of men as breadwinners and women as caretakers are not limited to less economically developed societies: despite the rapid rise in the United States of female labor force participation over the 20th century, greater than 70% of U.S. men and women continue to believe that a male should be the primary breadwinner of a household.[2] This research is motivated by the question: How are the returns to employment, both for oneself and others in the household, mediated by the gender of the employed? This question is of course dependent on context, and we describe ours of the Rohingya refugee camps in Bangladesh below. We then examine the impact of each on four bodies of outcomes by gender: wellbeing (psychosocial, physical, and cognitive), and how one perceives or engages with oneself, one's partner, and the outside world. To estimate this relationship, we engage individuals in the Rohingya refugee camps in Bangladesh in a randomized field experiment. Our setting is one of scarce employment opportunities: only 7.6% of refugees in our sample found any wage labor in the previous month. Our experiment implements three interventions. Relative to a control arm, in which individuals with survey participation, we design an employment arm, in which we offer surveying tasks for an average of four days per week for six weeks; a no employment arm, in which we offer no tasks; and a volunteer arm, in which we offer an equivalent opportunity. A comparison of the" work" arm to the control arm yields the value of employment in its entirety. A comparison of the work arm yields a causal estimate of the non-pecuniary value of employment. A comparison of the work arm to the volunteer arm enables a distinction between the psychosocial value of fulfilling the role of the 'breadwinner' of a family (to that of the experience of working). We enroll 2520 married couples into the experiment. While we offer a randomized treatment to only one member (male or female) of each couple, we survey both members, allowing us to document spillovers of each intervention on the spouses of treated individuals. This study makes three contributions. First, we estimate the **psychosocial value of employment by gender**. We distinguish between pecuniary and non-pecuniary channels through which employment may improve psychosocial wellbeing, as this is a central dimension along which social protection policies are designed. Were the entire value of employment derived from its pecuniary benefit, unemployment insurance could be a perfect substitute for employment. However, results from a recent pilot in which a volunteer arm and employment arm were tested in the same refugee context suggestively reflect asymmetry by

gender: among males, depression severity is equally reduced by employment or volunteering; among women, however, employment reduces depression severity while volunteering appears to, if anything, raise it. While only a pilot study of 300 households, these patterns suggest that, in this context, men derive psychosocial value primarily from the experience of working, while women derive value primarily from greater financial resources, finding consistent with a small body of longitudinal studies. Our present study engages a sample that is sufficiently powered to identify gendered differences in psychosocial outcomes. Second, we explore the channels through which these psychosocial benefits transpire, organizing these 'mechanisms' into how employment impacts one's perception of oneself, one's engagement with their domestic partner, and one's relationship with the outside world.

Perception of the self encompasses measures of purpose and self-worth vis-a˜-vis one's family and one's community as well as measures of how participation in housework or gainful employment align with one's gender identity. Engagement with one's partner includes revealed and stated measures of household bargaining power and patterns of intra-household decision-making. Relationship with the outside world is captured through measures of the quantity, length, and tenor of conversations that participants have with others in their community. We qualify this work with two important caveats. The features of employment that we vary in our design do not encompass all the channels through which employment might impact wellbeing. Numerous features of work - its sociability, physicality, or public nature, for example - may be equally relevant sources of value or harm. We choose to hold these features constant between our work and volunteer arm, and instead design our survey instruments to gather suggestive evidence around the relevance of these channels. But they remain beyond the experimental scope of this study and important questions for future research. In so far as we care about improving psychosocial wellbeing or increasing female empowerment, our design allows us to identify the optimal policy tool for doing so. For example, unemployment insurance may be an insufficient or potentially harmful 'replacement' for employment among men in contexts where the non-pecuniary benefits of employment. Notably, social protection policies such as job training programs, or work programs commonly determine the allocation of funding and the targeting of beneficiaries with a motivation of 'gender inclusion,' despite scarce evidence of the means through which male and female beneficiaries may derive value from the intervention: the WFP, for example, mandates a 60% female participation in their programs. We hope the findings of this study can offer clarity to the design of such programs as we probe the underlying value of employment to individual and household wellbeing among vulnerable and low-resource populations.

2. Research Context

2.1 Recent Events

In August of 2017, the Myanmar military executed a series of "Clearance Operations" in Rakhine State, Myanmar. The operations were targeted the Rohingya ethnic minority, who

have been denied citizenship in Myanmar since 1982 and now the world's largest stateless population. Over the course of four months, gang rapes and sexual violence were perpetrated against an estimated 18,000 women and girls, an estimated 36,000 Rohingya were thrown into fires, and at least 25,000 Rohingya were killed. Among those who survived, over 750,000 entered Bangladesh, building and settling into what is now the largest refugee camp in the world. They joined several hundred thousand Rohingya refugees from earlier episodes of ethnic violence, with the current population in the camps exceeding 900,000 individuals. This ignores the potentially positive long-term impacts of female labor force participation on social norms within the household and community, a necessary subject of future work. The goal of the present paper is to estimate whether backlash exists, relevant to the immediate welfare of female beneficiaries.

They joined several hundred thousand Rohingya refugees from earlier episodes of ethnic violence, with the current population in the camps exceeding 900,000 individuals[3].

2.2 Camp Life: Descriptive Statistics

There are presently 34 camps in Bangladesh, each subdivided into blocks ranging in population density from 60 to 130 households. These camps stretch from TEKNAF, a city on the southern tip of the country, north to UKHIYA. Each block within each camp is represented by a local Rohingya leader, or MAJHI, who is responsible for organizing distribution efforts and serving as a liaison between humanitarian organizations and the refugee community.

Because refugees are not legally allowed to work, many remain unoccupied in the camp. Among the scarce employment opportunities available are day laboring in agriculture or construction, operating street stalls, private tutoring, or assisting in NGO's women's cooking centers, child-friendly spaces, or health clinics. The majority of jobs available in the camps are provided by NGOs, several of whom have organized cash-for-work programs in the camp. Outside of the camps, a comparable population of Bangladeshis are likewise occupied in agriculture, fisheries, transit, or small street-side enterprises.

We describe our participants by gender. The average female refugee in our study is 28 years old, with 15.4% having zero formal education when in Myanmar. About 21.4% of females qualify as at least moderately depressed according to the PHQ-9 screening tool. A typical day in the life of a female, as captured through a recollection of time spent in the previous day on a variety of activities, consists of 8.4 hours sleeping, 0.4 hours engaged in wage work, 1.0 hours engaged in self-employed income-generating work, 2.8 hours doing chores outside the house, 3.0 hours doing chores inside the house, 2.8 hours actively taking care of elders, children, and the sick, 2.0 hours actively taking care of oneself, and 3.6 hours resting, relaxing,

---

[3] This ignores the potentially positive long-term impacts of female labor force participation on social norms within the household and community, a necessary subject of future work. The goal of the present paper is to estimate whether backlash exists, relevant to the immediate welfare of female beneficiaries.

or in religious activities. The average male refugee in our study is 31 years old, with 22% having zero formal education when in Myanmar. 18% qualify as at least moderately depressed. A typical day for a male participant consists of 8.7 hours sleeping, 0.8 hours engaged in wage work, 0.9 hours engaged in self-employed income-generating work, 2.7 hours doing chores outside the house, 1.7 hours doing chores inside the house, 2.3 hours actively taking care of elders, children, and the sick, 2.0 hours actively taking care of oneself, and 5.0 hours resting, relaxing, or in religious activities.

Table 1: Summary statistics, treated men

	(1)	(2)
	Mean	Std. Dev.
Age	31.38	6.37
People in HH	5.39	1.72
Work Last 30 Days	0.13	0.34
Hrs Idle/Day	4.02	2.54
Savings (BDT)	435.31	1059.44
Borrowing (BDT)	4496.12	11501.77
Consumption 2 Wks (BDT)	3785.23	3586.32
Depressed	0.18	0.38
Healthy Days (last 30)	26.45	3.81
Observations	1259	

Table 2: Summary statistics, treated women

	(1)	(2)
	Mean	Std. Dev.
Age	28.07	5.41
People in HH	5.46	1.76
Work Last 30 Days	0.00	0.06
Hrs Idle/Day	3.79	3.97
Savings (BDT)	200.84	725.30
Borrowing (BDT)	3952.29	29500.52
Consumption 2 Wks (BDT)	3441.67	3157.99
Depressed	0.21	0.41
Healthy Days (last 30)	23.92	5.39
Observations	1254	

3. Experimental Design

Sampling Strategy

The research team enlisted 2520 households across 10 camps (2E, 2W, 3, 4, 5, 7, 10, 17, 18, 19), with each camp apportioned into 4 to 7 blocks, and 14 to 42 sub-blocks, with the latter

as our unit of randomization. We selected nine households per sub-block. These households were recruited first through a random walk, in which the field team began near the center of a given sub-block, moved in a randomly selected direction, and proceeded to knock on doors and inquire about interest in participating in our study. If households voiced interest, the field team confirmed that there existed a married couple in the household who satisfied the following eligibility criteria: both members were between the ages of 18 and 45, able and willing to work, and had not worked for more than 10 hours in the past two weeks[4]. Prior to all field work, the research team secured permission from government authorities to operate in the camps.

Experimental Design

Our sample households spanned 280 sub-blocks, with each sub- block randomly assigned to one of the three experimental arms (80 blocks each) and a control group (40 blocks). Each household we selected was randomized into having the husband or the wife receives the respective intervention. All households were informed that the study would last six weeks, with surveyors returning weekly to a pre-assigned meeting point to conduct five-minute surveys. The randomized nature of treatment allocation was made public to all participants, with surveyors displaying the participant's randomized treatment status on their tablet screen.

Table 3: Treated beneficiaries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control	Work	no work	Voluntary	(1) vs. (2)	(1) vs. (3)	(1) vs. (4)
Age	29.69	29.32	30.1	29.78	0.37	0.36	0.67
People in HH	5.4	5.39	5.45	5.44	0.64	0.98	0.89
Math Literacy Index	3.23	3.27	3.18	3.23	0.85	0.58	0.66
Digit Span Index	6.07	6.11	6	6.22	0.79	0.31	0.63
Life Satisfaction Index	16.7	16.4	16.11	16.48	0.42	0.09	0.71
Self-worth Index	14.49	14.71	14.89	14.58	0.38	0.34	0.34
Purpose Index	14.48	14.58	14.84	14.53	0.94	0.86	0.93
Work Last 30 Days	0.08	0.07	0.07	0.05	0.69	0.78	0.28
Worked Myanmar	0.33	0.32	0.33	0.34	0.43	0.73	0.87
Hrs Idle/Day	3.94	3.85	3.98	3.87	0.65	0.34	0.86
Locus of Control	15.04	15.01	14.9	15.03	0.8	0.56	0.69
Health Days	24.84	25.17	25.32	25.24	0.28	0.12	0.51
PHQ	6.56	6.43	6.38	6.46	0.66	0.62	0.66
Stress Index	5.56	5.7	5.48	5.68	0.98	0.5	0.9
Diff. Conversations	12.56	12.63	12.55	12.29	0.26	0.84	0.22
Savings (BDT)	276.73	379.43	251.44	345.63	0.29	0.81	0.47
Consumption 2 Wks (BDT)	3710.52	3459.5	3645.39	3687.83	0.23	0.73	0.58
IPV Verbal Index	18.08	17.91	17.92	17.71	0.23	0.25	0.45
IPV Norms Index	6.75	6.97	6.99	6.87	0.65	0.13	0.76
Men in Home Norms Index	4.66	4.62	4.51	4.51	0.72	0	0.02
Women at Work Norms Index	0	0.01	0.04	0.05	0.26	0.82	0.3
Observations	360	717	720	716			
Joint F-Test					0.19	0.84	0.46

---

[4] We also confirmed that they were part of the most recent influx and were not a member of the *majhi's* family

Employment intervention details

Employees were asked to participate in a data collection task in which they selected fifteen of their same-sex neighbors and tracked these individuals' activities four times per day on a series of visual time-use worksheets (Figure 3). These neighbors remained anonymous to the surveyors and research team; we informed participants that we merely wished to get a sense of what the typical man or woman in their neighborhood did during a day in the camps. Workers were assigned three to five workdays per week for a total of 24 days of work over the six weeks, with all dates predetermined and noted on a calendar which was given as a gift to all participants. Worksheets were dropped off daily in a tamper-proof box at the home of a pre-assigned refugee neighbor within each sub-block (the 'facilitator').[5] At the end of each week, all workers convened at the facilitator's home, where an enumerator checked the participant's work for any mistakes (e.g. no missing sheets, submission made on the correct days; fifteen tick marks per sheet; no replication across days or obvious variation in handwriting suggesting someone else had done their work). However, participants were never at risk of losing their job nor otherwise being punished by the employer (the enumerator).

This work task was designed with several constraints in mind. First, because we wished to have equal parts male and female employment, we designed a task that did not require back-breaking manual labor. It did however require meaningful physical and mental exertion: workers engaged in repetitive movement throughout the day in the outdoor heat and muddy and cramped camp pathways, and completion of the task correctly required mental effort and attention. It was also designed to occupy a meaningful amount of time throughout the day, but was not a full-time occupation: the task was estimated to require approximately three hours total per work day; we collect this data and will report actual time investment upon study completion.

Second, because not all participants were literate, we designed a task that required no literacy or numeracy skills beyond counting. The time-sheet was a visual tool, composed of an exhaustive list of the activities that a man or woman in the camps may be engaged in. Workers needed only to place tick-marks below whichever activity they found their neighbors doing. Third, we designed a work task that would require workers to step outside of their home and see others, but did not require explicit socialization: workers could silently observe neighbors and complete their work sheets, or they could approach them and converse.

In sum, these considerations resulted in a work task that was similar to non-manual labor work that others in the camp may have access to (nearly all of which is supplied as "paid

---

5 Facilitators were also workers and had no access to the contents of the box. They were asked to place a sheet at the end of the day's work with the day's date, so that any submission 'below' that sheet was time-stamped to having been submitted on that day.

volunteer opportunities" by NGOs), accommodated the constraints of our study population, and was intended to be experienced as neither 'too good' nor 'too bad' in the setting of the refugee camps.

4. Empirical Strategy

Our primary objective is to understand how employment impacts wellbeing differentially across sex, and what mechanisms motivate these differences in the value of employment. We focus first on a series of wellbeing measures. We then pursue mechanisms. Employment can shift one's lived experience along multiple ins, altering how one perceives or engages with oneself, one's partner, and one's surroundings. We organize our outcomes accordingly, exploring patterns separately for men and women. The specific survey questions we ask to capture these outcomes, and the rational for including them, are detailed in Section 5.

MAIN EFFECTS	SPECIFIED OUTCOMES
Wellbeing: how work impacts how one feels	Psycho-social wellbeing, Physical wellbeing, Cognitive ability, Risk preferences

MECHANISMS	
Engagement with self: how work shapes perception of self	Self-Worth, Purpose, House norms (men), Work norms (women)
Engagement with partner: how work shapes household interactions	Intra-household Decision-Making, Bargaining (Stated, Revealed)
Engagement with outside world: how work shapes relationship with the outside world	Sociability

Employment status can also affect one's spouse, which may have direct implications for own wellbeing as well as the total welfare of the household. We therefore also track how own treatment status impacts a partner's wellbeing and the partner's engagement with themselves, their spouse, and others. These underlying mechanisms are necessarily interconnected, and our exploration of mechanisms is not intended to establish causal links between these channels. The objective of this latter series of exercises is rather to underscore which psychosocial features shift due to an employment treatment, and how these may differ depending on the sex of the employed.

4.1 Main effects

We first capture the impact of work on our measures of wellbeing with the following specifications:

$$Y_{ibc} = \beta_0 + \vartheta_1 Work_{ibc} + \vartheta_2 \textbf{No work}_{ibc} + \vartheta_3 Volunteer_{ibc} + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

$$Y_{ibc} = \beta_0 + \beta_1 Women + \beta_2 Work_{ibc} + \beta_3 Work_{ibc} * Women + \beta_4 No\ work_{ibc} + \beta_5 No\ work_{ibc} * Women + \beta_6 Volunteer_{ibc} + \beta_7 Volunteer_{ibc} * Women$$
$$\gamma c + \delta e + Xibc + \varepsilon ibc$$

where Y_{ibc} represents our measures of wellbeing (psychosocial wellbeing, physical wellbeing, cognitive ability, and risk preferences) for individual i in block b and camp c. X_{ibc} is a vector of sociodemographic controls selected via double-selection LASSO to maximize precision, and ϵ_{ibc} is an error term clustered at the block level. We include fixed effects for camp γ_c and enumerator δ_e.[6] The first specification pools men and women, while the second includes the interactions of treatment with sex to identify how impacts differ by sex.

Test 1 If work generates positive psychosocial benefits, then $\vartheta_1 \geq 0$.

Test 2 If work delivers psychosocial benefits equal to or beyond the benefits of income alone, then $\vartheta_1 \geq \vartheta_2$. If work delivers psychosocial benefits equal to or beyond the benefits of doing something productive alone, then $\vartheta_1 \geq \vartheta_3$.

Test 3 If gainful employment impacts men and women differently, then $\beta_3 \neq 0$. The sign of β_3 remains *ex-ante* ambiguous.

Test 4 The benefits of work relative to those of productive activities may be different for men and women as well.

4.2 Mechanisms

Conditional on finding that the experience of working itself is of psychosocial value, we then consider what about this experience might generate this impact. We organize our channels into how employment impacts one's perception of oneself, one's engagement with their domestic partner, and one's relationship with the outside world. We set each as an outcome variable to our main specification, presenting results separately for men and women. Our objective is to peel back the various layers of experience that employment and volunteering may impact, something which is likely to be distinct by sex. Given the wide and inconclusive literature on these relationships (e.g. impact of employment on male / female bargaining

---

6 We follow and include enumerator fixed effects to account for the fact that respondents' answers may be influenced by the way enumerators ask more sensitive questions.

power), we do not specify a test of difference by sex along these outcome measures. But we do find an analysis run separately by sex to be informative of why these treatments may have different implications by sex and therefore what features of employment, or volunteering a policymaker may wish to focus on to maximize psychosocial wellbeing.

4.2.1 Dimensions of one's own psychosocial wellbeing

We first run the following specification for men and women separately:

$$Y_{ibc} = \beta_0 + \vartheta_1 Work_{ibc} + \vartheta_2 \textbf{\textit{No work}}_{ibc} + \vartheta_3 Volunteer_{ibc} + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

where Y_{ibc} represents the set of outcomes pertaining to mechanisms (engagement with self, engagement with partner, and engagement with the outside world), which we detail in Section 5 for individual i in block b and camp c, and γ_c, δ_e, X_{ibc}, ϵ_{ibct} are as defined above.

Test 6 If work generates positive impacts on our mechanisms of interest, we expect to find that $\vartheta_1 \geq 0$.

4.2.2 Dimensions of one's partner's psychosocial wellbeing

The psychosocial value of employment is likely to be informed not only by one's own direct experience but also by how those nearest you behave in response. For example, if men who are otherwise unemployed feel threatened by an employed wife, any resulting depression or aggression may reduce the women's psychosocial wellbeing. Conversely, if wives of formerly unemployed men become happier when their spouse resumes the position of the breadwinner, then men's self-worth and psychosocial wellbeing may be amplified. We therefore also apply our main specification to the spouse's wellbeing:

$$Y_{jbc} = \beta_0 + \vartheta_1 Work_{ibc} + \vartheta_2 \textbf{\textit{Nowork}}_{ibc} + \vartheta_3 Volunteer_{ibc} + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

where Y_{jbc} represents the set of outcomes outlined in Section 5 for individual i's partner j in block b and camp c, and γ_c, δ_e, X_{ibc}, ϵ_{ibct} are as defined above.

Test 7 If work generates positive or negative impacts for the spouse on our mechanisms of interest, $\vartheta_1 \neq 0$.

Beyond the partner's instrumental role in one's own wellbeing, we are also interested in the partner's wellbeing as an outcome in its own right. Employment is likely to impact households as a unit, and a complete picture of their psychosocial implications requires measuring, at the least, their impact on both heads of a household. It remains an empirical question whether women or men working is more valuable to total household welfare; in settings with limited resources, identifying who in a household should be targeted for employment to maximize welfare is a valuable policy objective.

4.3 Willingness to work

While the preceding exercises estimate the value of employment on our participants' wellbeing, they do not provide evidence of whether the employed are aware of these gains - a question relevant to the sustainability of employment or volunteering programs. We therefore also engage in an incentivized revealed preference exercise to estimate whether and to what extent households value the opportunity to be employed. Having experienced the work task and therefore able to realistically value the work, at endline, we surprise individuals in the employment and volunteer arms with the opportunity to engage in an additional week of work.

For those who are not willing to work, we use the incentivized Becker- DeGroot-Marschke (BDM) method. We use the 13 Kolmogorov-Smirnov tests to determine whether these labor supply curves differ significantly between the sexes.

4.4 Robustness

In this case, the psychosocial benefits we identify may stem from expectations of future work rather than the act of engaging in work for pay during the study period. To limit such expectation formation, we emphasize repeatedly throughout the study that the work opportunity we provide will only continue for six weeks. Additionally, we also embed experimental variation in the salience of 'work experience,' the channel through which we suspect expectations of future work to flow. Specifically, we present certificates of participation to a randomized half of our sample (across all treatment groups). These certificates provide documentation of the beneficiaries' involvement with our project, and are intended to provide an explicit boost to their resumes.[7] The effect of the embedded certificate randomization is estimated via the following regression:

$$Y_{ibc} = \beta_0 + \beta_1 Certificate + \beta_2 Work_{ibc} + \beta_3 Work_{ibc} * Certificate + \beta_4 Large\mathbf{work}_{ibc} + \beta_5 Large\mathbf{work}_{ibc} * Certificate + \beta_6 Volunteer_{ibc} + \beta_7 Volunteer_{ibc} * Certificate + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

where Y_{ibc} represents our measure of psycho-social wellbeing for individual i in block b and camp number c, and γ_c, δ_e, X_{ibc}, ϵ_{ibct} are as defined above.

Test 9 We are interested in whether $\beta_3 \geq \beta_5 \geq 0$: in other words, whether there is any differential impact of providing the certificate on psychosocial outcomes, and whether this differential impact is over and above that of reciprocity alone (which is identified from β_5: if individuals experience a warm glow from receiving the certificate, this will be captured by a positive effect among those in the treatment arm).

---

7 The certificate read "I engaged with RTM International to do data collection". It was written this way in order to be generic enough to apply to all the individuals in the experiment, all of whom were providing us data from the weekly surveys.

While this experiment provides some evidence for or against the role of future work expectations, it is not definitive, as the certificate may not sufficiently raise the perceived probability of future work.

5. Data Collection and Survey Instruments

We describe here our process for data collection and details on the key set of outcome variables that populate our survey instruments.

5.1 Timeline

Survey teams entered each block to recruit the study sample and conduct the baseline survey. Each household was informed that we may have an opportunity for them to work for up to four hours per day for four days a week over six weeks. We clarified that we had not yet secured our funds for this activity, and we would not have enough work opportunities for everyone, but we wished to know whether both members of the married couple would be able and interested in working for us, and whether they would be willing to meet with us for ten minutes every week for six weeks to answer survey questions in the case that we could not offer them work. For those who consented, we proceeded with the baseline survey, which was administered to both members of the couple. The field team then revisited these households the following week to reveal which treatment status they had been randomized to and conduct the first midline survey. Thereafter, brief surveys were conducted each week, followed by payment disbursal. The endline survey took place five weeks after the start of work, allowing us to capture the effects of working while those who were employed or in the volunteer arm were still engaged in the activity. We conducted a final follow-up survey approximately six weeks after endline, wherein we collected basic psychosocial measures to ensure that the halting of the interventions did not negatively impact our treated participants. We also used this opportunity to surprise former work and volunteer treatment groups with one final week of work in order to estimate participants' valuation of work.

5.2 Outcome Variables

All outcomes we describe below are collected via the surveys described above. The questions in these surveys were drawn from previous work in the camps and piloted extensively with households that were not included in the study sample. We categorize our outcomes into main effects and mechanisms. We organize the underlying channels navigating this relationship into three broad buckets: how work shapes the perception of self, household interactions, and relationships with the outside world.

Given the wide range of hypotheses we test, we account for multiple hypotheses by computing False Discovery Rate (FDR) q-values. The primary groups of outcomes are listed below, and for each of these outcomes we will construct indices (where possible using inverse covariance weighting) and report both p-values and sharpened q-values.

5.2.1 Main effects

We measure five dimensions of **psychosocial wellbeing**. First, we measure depression using a nine-question depression scale of the Patient Health Questionnaire (PHQ-9), a standardized screening tool that assesses mental and emotional health disorders. Second, we adapt Levenson's Multidimensional Internal Locus of Control Scales, which asks four questions about the degree to which people believe that they, as opposed to external forces, have control over the outcomes in their lives. Third, we measure life satisfaction with Diener's Satisfaction with Life Scale (including four out of the five statements). Fourth, we measure stress with three questions inspired by Cohen's Perceived Stress Scale. Fifth, we assess respondents' sense of stability by asking them how secure they feel at the moment and expect to feel in the future.

We supplement psychosocial wellbeing with three other measures of wellbeing. We measure **physical wellbeing** by asking respondents how many days they were sick in the past month. We measure **cognitive ability** with a digit-span memory test using both forward and backward sequences of numbers and four of basic arithmetic problems. Finally, we capture respondents' **risk preferences**. We ask respondents to make ten choices between a lottery or a certain amount, where the lottery has increasingly favorable odds. We then randomly draw once choice for their prize.

5.2.2 Main Mechanisms

Engagement of self: Many individuals develop a sense of identity through work they do. We capture how our interventions may alter respondents' sense of self in four ways. First, we construct an index of **purpose** by asking respondents to first consider the person in their community (family) who contributes the most to their respective community (family), and then have the respondent rank where they would place themselves relative to this individual.

Second, we construct an index of **self-worth** by having respondents consider the person in their community (family) who they respect the most in their community (family), and then have the respondent rank where they would place themselves relative to that individual.

Third, we construct an index of **norms for men within the household** from two questions. Specifically, we ask them to agree/disagree with the following statement "A husband who helps his wife with the household chores should not be respected" and "A husband who makes important decisions jointly with his wife is weak".

Finally, we construct an index of **norms for women in the workplace** composed of four questions. First, we ask how many days (and hours per day) respondents think that women should be allowed to work outside the home (but inside the block), and outside the home (and outside the block). Next, we ask respondents to agree/disagree with the following statement: "A wife who prioritizes work outside the home over household-chores is not a

good wife". Finally, we adapt a survey module that asks respondents to choose a hypothetical husband (wife) for a daughter (son) they have (or might have in the future). The respondent must choose between two men with the same education and income but only one of them would allow the respondent's daughter to work outside for pay, and between two women with the same education and the same income, but one of the women wishes to work outside the home for pay.

Beyond the above, we also collect but do not pre-specify analyses for a measure of women's (men's) perception of an ideal women (man) and respondents' aspirations. We measure the former by adapting a survey module that asks respondents what features they place the greatest importance on for themselves (e.g. financial stability; good work ethic; taking good care of children and other family members; being admired and respected by people in the community; putting others needs before their own). We estimate respondents' aspirations by eliciting a set of goals that the respondent may have and asking how actively they are trying to reach their goals.

Engagement with partner: Employment may also affect how individuals engage with their partners. We capture a respondent's relationship with his/her partner in four ways. First, we include a collection of **household bargaining** measures.

We play an incentivized bargaining game. We invite both members of a couple to decide how to allocate 250 takas between themselves and their spouse. Both respondents play this game independently with separate enumerators. Their decisions are entirely private. The enumerator records the amount the respondent allocated to themselves on a chit, and records the amount the respondent's spouse allocated to their respondent on a separate chit (information they collect from the other enumerator who is playing the game in parallel with the spouse). We then ask the spouses to sit together and come to an agreement over how to divide the 250 takas between each other, and these allocations are also transcribed onto chits. The enumerator then places four chits in a tin: the amount their respondent allocated to themselves privately, the amount the respondent's spouse allocated to the respondent, the amount the respondent and the spouse allocated to the respondent in the joint-round, and a random amount. The random chit is designed to ensure that the respondent cannot infer which amount their spouse put in. The respondent is then instructed to pick one of the four chits, and this is the amount of gift that they receive.

Beyond this revealed preference measure of bargaining power, we also ask respondents how they engage in conversation with their partner through a series of questions. Specifically, we ask each respondent how often their spouse considers their opinion, whether they share their opinion with their spouse when they disagree and try to change their mind, how often or willing their spouse is to change their mind in the event of a disagreement, and who makes the final decision in a disagreement. Finally, because household bargaining norms may be

deeply embedded and difficult to change in practice, we ask one final question to capture respondents' beliefs about women's bargaining power: we ask respondents to agree or disagree to the statement "A wife who frequently expresses her opinion in the household is overbearing/talks too much."

Third, we measure **intra-household decision making** by asking a series of nine questions designed to track how decisions over consumption are made within the household. We loop over the following questions for five consumption items (small household purchases, large household purchases, child, health, luxury).

We ask: 1) how much the household spends on a particular item per month, 2) who makes the purchase, 3) who the respondent *thinks* should make the purchase, 4) who provides the money for the purchase, 5) who the respondent *thinks* should provide the money for the purchase, 6) who decides how much to spend on the item, 7) who the respondent *thinks* should decide how much to spend on the item, 8) how much money the respondent can spend on the item before consulting their partner, 9) how much money respondents *think* they should be able to spend on the item before consulting their partner.

We standardize each of these eight measures and combine (6) and (8) into an index that measures consumption decision-making. We combine (7) and (9) into an index that measures beliefs around consumption decision making. We explore the other measures but do not pre-specify them.

Relatedly, we ask a series of four questions designed to track how decisions over time-use are made within the household. We loop over the following questions for four-time actions (raising children, working outside the home, indoor household chores, and outdoor household chores). We ask: 1) who in the household performs the task, 2) who the respondent thinks should perform the task, 3) who in the household decides who performs the task, 4) and who the respondent thinks should decide who perform the task. As above, we standardize each of these measures and focus on (3) to measure time use decision making. We also focus on (4) to measure beliefs around time-use decision making. We explore the other measures but do not pre-specify them.

Engagement with outside world:
Working often involves some degree of engagement with the outside world. In the context of our study, the work task we designed required participants to step outside their tents, potentially sparking more opportunities for conversations with others. We measure participants' **sociability** by asking respondents how many different people they had a conversation with yesterday, how many of these conversations were longer than fifteen minutes, and how many of these conversations made them feel happy.

5.2.3 Additional Mechanisms

While we include questions to track consumption, time-use, and skills, our previous work suggests that these are not the operative margin through which work shifts wellbeing. We describe these outcomes below, but do not pre-specify them.

Financial portfolio We capture loans by asking respondents how much money households have borrowed, and how much they have lent to others. We also ask whether households would be able to cover an unexpected expense (1000 taka) if needed this week. We measure consumption by asking respondents how much money they spent on the following categories in the past two weeks: daily groceries (e.g. rice, lentils, oil), less-accessible consumables (e.g. meat, fish, fruit, vegetables), luxury items (e.g. cigarettes, tea and coffee, tobacco), education, healthcare, loan repayments, lending to others, transportation, festivals, bribes/extortion, small non-food household items (e.g. phone bill, mosquito nets, kitchen materials), and large household expenditures (home improvement, furniture).

Time-use Commitment to an employment activity may alter the set of activities one otherwise engages in during a typical day. We rely on a new time-use survey module. The paper demonstrates the benefits of this new approach: it is low cost and accurate in capturing individuals' average time use. The exercise invites respondents to think about the time they spent on a broad set of activities (e.g. sleeping, income generating activities, household chores, child care and leisure) in the past 24 hours. The enumerator has 24 chips, where each chip is equivalent to one hour. The enumerator converts respondents' narratives into these times use categories by allocating the chips across the major activities, which are represented by pictures.

Skills We also measure whether respondents' skills change as a result of engaging in work. We capture respondents' hard and soft-skills. For soft-skills, we ask the respondent to agree/disagree with the following statements: "When you meet a new person, you can speak to him/her easily," and "When you have a busier day than usual, you can finish all your work in time." For hard-skills, we provide instructions for drawing an image and ask respondents to repeat the instructions back to the enumerator and draw the image. The enumerator notes whether respondents are able to explain the instructions and rates the quality of the drawing.

5.3 Willingness to work

We apply the incentivized Becker-DeGroot-Marschke (BDM) method among work and volunteer group respondents in the follow-up survey (both of whom now have experience with the work task), and ask them if they are willing to complete an additional week of work at various wages. Pairing these responses with the number of days of additional work the respondent actually completes, we plot their labor supply curve.

5.3.1 Weekly Surveys

Every week, our team visits households to disburse payments and conduct a short survey that captures three outcomes of interest.

Main – Wellbeing We ask two questions about general wellbeing and three questions about stress. Respondents are asked to choose which among six faces best describe how they were feeling yesterday, and how many days they felt good mentally overall in the past week. We measure stress with three questions inspired by Cohen's Perceived Stress Scale.

Mechanism - Engagement with self We capture purposefulness by asking respondents to consider the person in their family who contributes the most to their family, and then ask the respondent to rank themselves relative to that individual.

Mechanism - Engagement with partner We adapt our baseline questions and compute an index by asking the main beneficiary (not the spouse) three questions: how many times in the last week did your partner say something that you disagreed with, how many times did you express your disagreement vocally, and how many times did your spouse's position change because of what you said?

5.3.2 Robustness

Work Expectations We emphasize repeatedly that our work program will end after six weeks. During our endline survey, we remind households that we are closing the program and we do not anticipate being able to continue. However, to capture respondents' expectations about future work opportunities, we ask how many days (and hours per day) they expect to be working in approximately two months' time.

6. Power

We present power calculations for two primary comparisons. We first compare primary treatments to one another, pooling between men and women (e.g. the work to volunteer treatment arms). We then compare male to female samples within a single treatment arm (e.g. women to men within the work group). For all calculations, we assume a 5% significance level and vary the inter-cluster correlation (ICC) between 0.01 to 0.2.[8]. Using these assumptions, we then calculate power across a range of standardized effect sizes. The power estimates shown below are conservative, as we do not account for the inclusion of baseline outcomes and lasso selected controls, which should improve the precision of the estimated treatment effect. For the full main treatment arm comparisons, the experiment is well powered to detect 0.15 standard deviation effect sizes at low ICC.

---

[8] In the baseline data, actual ICC ranges from 0.01 to 0.12 with most core outcomes at the lower end of this range.

At higher ICC levels, minimum detectable effect sizes range from 0.2-0.25 standard deviations. Observed ICCs at baseline tend towards the low end of the examined range while standardized effect sizes are typically between 0.20.25, suggesting that our primary specifications are well-powered.

Power Calculation for Main Effects

Effect Size	Inter-Cluster Correlation			
	0.01	**0.05**	**0.1**	**0.2**
0.15 Std Dev	0.78	0.67	0.56	0.42
0.20 Std Dev	0.95	0.89	0.81	0.65
0.25 Std Dev	1.00	0.98	0.94	0.84
0.30 Std Dev	1.00	1.00	0.99	0.94

Within each arm, the gender comparison is slightly less well-powered. At low ICC, we are powered to detected effect sizes between 0.2 and 0.25 standard deviations. At higher ICC, we are only powered to detect effect sizes at and above 0.25 standard deviations. That said, this test is still adequately powered considering the low observed ICC found in the data.

Power Calculation for Gender Comparison

Effect Size	Inter-Cluster Correlation			
	0.01	**0.05**	**0.1**	**0.2**
0.15 Std Dev	0.51	0.46	0.41	0.34
0.20 Std Dev	0.75	0.70	0.64	0.54
0.25 Std Dev	0.91	0.87	0.82	0.73
0.30 Std Dev	0.98	0.96	0.93	0.87

7. Experimental Results

7.1 Completion of work
We first establish that participants engaged in the treatment they were offered. Figure 1 exhibits the fraction of individuals assigned to the data-collection task who completed their work. Both groups are consistently above a 90% completion rate, indicative of participants' desire to engage in the work. Meanwhile, Table 9 shows that the Work and No work groups significantly increased their savings by over 500 BDT on average, also decreasing their outstanding borrowed amount by a similar amount.

7.2 Impact of employment
Table 10 presents the treatment coefficients for psychosocial outcomes. Relative to those in the control group, individuals in the Work arm experience a 0.09-unit improvement in their lives, and feel more stable in the present and the anticipated long-term future. In general, there is suggestive evidence that those offered employment exhibit greater life satisfaction, are more sociable, possess higher self-worth, feel greater control over events.

The employment arm not only improves psychosocial wellbeing relative to the control arm, but also yields significantly larger improvements in psychosocial wellbeing than the no work and voluntary arms. We can reject equality of effects between both respective groups and employment at the one percent level. This result is manifested particularly for the PHQ and stress sub components. In other words, what matters for the psychosocial value of work is gainful employment –the combination of work and getting appreciation. Productive activity and appreciation may be complements; not merely substitutes.

7.3 Nature of the worker
Treated women and men (table 11 and 12) exhibit comparable improvement in their psychosocial index, with Work significant at the one percent level for men, and at the ten percent level for women. However, the spillover effect of treatment on the partners is markedly different. Husbands of treated women exhibit no change in their psychosocial welfare, with suggestive evidence of a negative impact on those in the no work and Voluntary groups. On the other hand, wives of treated men appear to benefit even more than if they had received the treatment themselves. Wives in the Work arm experience a 0.1-unit improvement in their psychosocial index, 0.09 units in the No work group, and 0.07 units in the Voluntary group.

This evidence thus suggests that households in which men receive the employment opportunity exhibit greater total gains in wellbeing than those in which women receive the employment opportunity. However, we urge caution on drawing policy conclusions from this result. It may be the case that women experience longer-term benefits than men.

Table 4: Intervention Timeline by Weeks

T0	Baseline Survey
T1 (T0 +1 week)	Intervention + Midline 1
T2 (T0 +2 week)	Work Submission + Midline 2
T3 (T0 +3 week)	Work Submission + Midline 3
T4 (T0 +4 week)	Work Submission + Midline 4 + Certificate Delivery
T5 (T0 +5 week)	Work Submission + Endline
T6 (T0 +6 week)	Work Submission + Follow up

The baseline was launched in January 2023, and the study will conclude in July 2023. With approximately 50 enumerators working each day, we need to phase blocks of respondents into the study over time (in batches). The first batch of respondents were baselined January 23rd and their endline will be exactly 5 weeks from January 23. Similarly, the second batch of respondents were baseline on January 24th, and their endline will be exactly 5 weeks from January 24.

Table 5: Outcome Variable Description

Psychological Well-being

PHQ9	The standardized total score of 9 questions from the Patient Health Questionnaire-9 (PHQ9)
Locus of Control	The standardized total score from responses to four locus of control questions. "In the last 7 days, how many days did you feel that to a great extent your life is controlled by accidental/chance happenings..."
Life Satisfaction Index	A standardized average of survey responses to four questions from Diener's standardized scale, responses made along a six-point Likert scale.
Stress Index	The standardized total score from three elements of adapted from the Cohen Stress scale. "How many of the last 7 days have you [been able to fall asleep peacefully / felt nervous / felt frustrated]?"

Stability Index	The standardized total score from responses to two stability questions using a Cantril ladder. "How secure [do you feel / think you will feel] [at present / five years from now]"

Other Well-being

Physical Health	Number of days sick in the last 30 days.
Cognitive Ability	A standardized weighted index of the number of correct responses to i) a digit span (forward and backward) memory test and ii) basic arithmetic problems including addition, subtraction, multiplication, and division.
Risk Preference	Measured using incentivized responses to the multiple price list decisions adapted from Holt-Laury and Sprenger (2002).

Engagement of Self

Purpose	The standardized total score from the responses on a scale from 1 to 10 to two questions: "Think of a person you know who contributes the most in your [family / community]. If that person is a 10 where would you put yourself?"
Self-Worth	The standardized total score from the responses on a scale from 1 to 10 to two questions: "Think of a person you know who you respect the most in your [family /community]. If that person is a 10 where would you put yourself?"
Norms for household men within the Household	"A husband who helps his wife with the household chores should not be respected" and "A husband who makes important decisions jointly with his wife is weak", responses made along a four-point Likert scale.

Norms for women in the workplace	Hours per week that women should be allowed to work outside the home (but inside the block), and outside the home (and outside the block). Next, "A wife who prioritizes work outside the home over household-chores is not a good wife", responses made along a four point Likert scale. Finally, respondents choose a hypothetical husband for a daughter between two men with the same education and income, but only one of them would allow the respondent's daughter to work outside for pay. For a son, they choose between two women with the same education and the same income, but one of the women wishes to work outside the home for pay.
Engagement with partner	
Household bargaining	An incentivized bargaining game where both members of a couple decide how to allocate 250 taka. Also, five questions about how often their spouse considers their opinion and what happens in the event of a disagreement, responses made from five frequency options. Finally, "A wife who frequently expresses her opinion in the household is overbearing/talks too much", responses made along a 4-point Likert scale.
Intra-household decision making	A standardized index from responses to nine questions about consumption decisions and norms, looped over five categories (small household purchases, large household purchases, child, health, luxury). Also, a standardized index from responses to four questions about time use, looped over four categories (raising children, working

outside the home, indoor household chores, and outdoor household chores).

Engagement with outside world

Sociability (Total)	The total number of conversations in the past day with adults.
Sociability (Positive)	The total number of conversations in the past day with adults that the respondent felt were positive.
Sociability (Duration)	The total number of conversations that were longer than 15 minutes.

Other Mechanisms (not pre-specified)

Financial portfolio	Amount in savings total and over 90 days. Also, how much currently lent and borrowed to others, and whether respondent could cover an unexpected 1000taka expense. Finally, consumption across 12 categories in the past two weeks.
Time-use	Amount of time spent on a set of eight activities in the past 24 hours.
Skills	Agree or disagree to: "When you meet a new person, you can speak to him/her easily" and "When you have a busier day than usual, you can finish all your work in time". Also, ability to follow instructions from an enumerator to draw an image.

Table 6: Outcome Variable Collection Periods

Psychological Well-being	Baseline	Weekly	Endline	Follow-up
PHQ9	X		X	
Locus of Control	X		X	
Life Satisfaction Index	X		X	X
Stress Index	X	X	X	
Stability Index	X		X	
General Wellbeing		X	X	
Other Well-Being				
Physical Health	X		X	
Cognitive Ability	X		X	
Risk Preference	X		X	
Engagement of Self				
Purpose	X		X	
Self-Worth	X		X	
Norms for Men	X	X	X	X
Norms for Women	X		X	X
Engagement with partner				
Household bargaining	X		X	
Intimate partner violence	X	X*	X	X
Intra-household decision-making	X		X	X
Engagement with outside world				
Sociability (Total)	X		X	
Sociability (Positive)	X		X	
Sociability (Duration)	X		X	
Other Mechanisms				
Financial portfolio	X		X	
Time-use	X		X	
Skills	X		X	

Table 7: Financial

	(1) Savings	(2) Borrowing	(3) Lent	(4) Spend 1000	(5) Total Consumption
	Panel: All Treated				
Work	557.336***	-664.135***	0.001	0.114***	-54.438
	(81.575)	(254.619)	(0.014)	(0.024)	(268.091)
No work	519.004***	-508.177*	-0.010	0.110***	73.310
	(82.631)	(271.095)	(0.013)	(0.025)	(275.769)
Voluntary	20.253	89.513	-0.019	0.006	9.337
	(81.096)	(257.441)	(0.012)	(0.027)	(283.117)
Control Mean	375.913	2479.31	0.052	0.739	4140.761
Shrp. q-val Work	0.001	0.010	0.613	0.001	0.613
Test Work=No work	0.555	0.445	0.366	0.806	0.505
Test Work=Unpaid	0.000	0.000	0.068	0.000	0.770
Observations	2264	2424	2426	2426	2429

Table 8: Psychosocial value of employment versus no work and voluntary work

	(1) Ps Index	(2) PHQ	(3) Stress	(4) Life Sat.	(5) Social	(6) Purpose	(7) Self-Worth	(8) Control	(9) Stability
	Individual Components of PS Index								
Work	0.089***	0.164**	0.099	0.099**	0.060	0.093	0.088	0.089	0.060
	(0.031)	(0.065)	(0.063)	(0.048)	(0.044)	(0.057)	(0.057)	(0.055)	(0.056)
No Work	0.028	0.022	-0.066	0.091**	-0.083*	0.081	0.093	0.006	0.006
	(0.032)	(0.065)	(0.062)	(0.045)	(0.046)	(0.057)	(0.057)	(0.053)	(0.056)
Voluntary	0.019	0.072	-0.042	-0.021	0.017	0.050	0.033	0.081	-0.059
	(0.032)	(0.065)	(0.060)	(0.046)	(0.050)	(0.055)	(0.053)	(0.055)	(0.057)
Control Mean	0.004	0.000	-0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000
Shrp. q-val Work		0.111	0.166	0.154	0.180	0.166	0.166	0.166	0.194
Test Work = No Work	0.007	0.010	0.005	0.840	0.000	0.791	0.911	0.069	0.233
Shrp. q-val Work= No Work		0.023	0.020	0.593	0.002	0.593	0.593	0.869	0.009
Test Work = Voluntary	0.002	0.082	0.015	0.003	0.289	0.313	0.176	0.869	0.009
Shrp. q-val Work= Voluntary		0.115	0.036	0.029	0.218	0.218	0.197	0.484	0.034
Observations	2429	2429	2425	2425	2425	2425	2425	2425	2425

Table 9: Psychosocial Wellbeing of Treated Women

	(1) Ps Index	(2) PHQ	(3) Stress	(4) Life Sat.	(5) Social	(6) Purpose	(7) Self-Worth	(8) Control	(9) Stability
Panel A: Treated Women	Individual Components of PS Index								
Work	0.081*	0.187**	0.107	0.113	-0.060	0.047	0.084	-0.016	0.128**
	(0.042)	(0.086)	(0.093)	(0.072)	(0.090)	(0.076)	(0.071)	(0.074)	(0.064)
No Work	0.041	0.035	-0.053	0.128**	-0.259**	0.023	0.065	0.012	0.052
	(0.041)	(0.084)	(0.092)	(0.064)	(0.089)	(0.072)	(0.067)	(0.076)	(0.069)
Voluntary	0.014	0.101	-0.027	-0.051	-0.053	0.021	0.034	0.055	0.006
	(0.040)	(0.086)	(0.092)	(0.066)	(0.110)	(0.074)	(0.064)	(0.079)	(0.068)
Control Mean	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
Shrp. q-val Work Test		0.225	0.429	0.301	0.621	0.621	0.429	0.667	0.225
Work = No Work	0.224	0.036	0.051	0.806	0.005	0.663	0.742	0.652	0.209
Test Work = Voluntary	0.035	0.263	0.119	0.012	0.941	0.662	0.377	0.292	0.054
Observations	1212	1212	1210	1210	1210	1210	1210	1210	1210

Table 10: Psychosocial Wellbeing of Partner Men

	Panel A: Partner Men			Individual Components of PS Index					
	(1) Ps Index	(2) PHQ	(3) Stress	(4) Life Sat.	(5) Social	(6) Purpose	(7) Self-Worth	(8) Control	(9) Stability
Work	-0.002	0.002	0.026	-0.034	-0.014	0.082	0.049	-0.050	0.002
	((0.034)	(0.083)	(0.088)	(0.071)	(0.074)	(0.087)	(0.102)	(0.073)	(0.097)
No Work	-0.039	-0.001	-0.076	-0.131**	-0.057	0.045	0.055	-0.058	-0.004
	(0.034)	(0.085)	(0.087)	(0.063)	(0.074)	(0.086)	(0.098)	(0.078)	(0.094)
Voluntary	-0.051	-0.034	-0.090	-0.091	-0.026	0.055	-0.034	-0.125*	-0.082
	(0.033)	(0.083)	(0.088)	(0.063)	(0.074)	(0.085)	(0.099)	(0.076)	(0.097)
Control Mean	0.053	-0.000	0.000	0.000	0.000	0.000	0.000	-0.000	0.000
Shrp. q-val Work Test		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Test Work = No Work	0.182	0.956	0.175	0.131	0.437	0.555	0.943	0.899	0.928
Test Work = Voluntary	0.068	0.539	0.101	0.352	0.837	0.668	0.307	0.197	0.270
Observations	1175	1175	1175	1175	1175	1175	1175	1175	1175

Table 11: Psychosocial Wellbeing of Treated Men

	Panel A: Partner Men			Individual Components of PS Index					
	(1) Ps Index	(2) PHQ	(3) Stress	(4) Life Sat.	(5) Social	(6) Purpose	(7) Self-Worth	(8) Control	(9) Stability
Work	0.106**	0.136*	0.083	0.091	0.134**	0.139*	0.095	0.200***	-0.016
	(0.036)	(0.081)	(0.084)	(0.066)	(0.062)	(0.079)	(0.086)	(0.072)	(0.078)
No Work	0.024	0.009	-0.087	0.059	-0.010	0.154*	0.129	0.001	-0.040
	(0.039)	(0.082)	(0.086)	(0.067)	(0.065)	(0.079)	(0.082)	(0.069)	(0.078)
Voluntary	0.021	0.025	-0.077	0.009	0.064	0.080	0.035	0.094	-0.121
	(0.040)	(0.083)	(0.086)	(0.065)	(0.066)	(0.079)	(0.084)	(0.071)	(0.080)
Control Mean	0.010	0.000	0.000	0.000	-0.000	0.000	-0.000	0.000	0.000
Shrp. q-val Work Test		0.159	0.228	0.223	0.122	0.159	0.223	0.047	0.456
Test Work = No Work	0.005	0.057	0.027	0.532	0.009	0.807	0.612	0.001	0.710
Test Work = Voluntary	0.004	0.098	0.029	0.090	0.195	0.343	0.382	0.101	0.108
Observations	1217	1217	1215	1215	1215	1215	1215	1215	1215

Table 12: Psychosocial wellbeing of partner women

	Panel B: Partner Women			Individual Components of PS Index					
	(1) PS Index	(2) PHQ	(3) Stress	(4) Life Sat.	(5) Social	(6) Purpose	(7) Self-Worth	(8) Control	(9) Stability
Work	0.102***	0.213***	0.148	0.034	-0.101	0.170***	0.248***	-0.054	0.120
	(0.044)	(0.100)	(0.097)	(0.074)	(0.100)	(0.062)	(0.069)	(0.076)	(0.088)
No work	0.094***	0.058	0.052	0.166***	-0.091	0.044	0.175***	0.014	0.101
	(0.045)	(0.097)	(0.094)	(0.071)	(0.104)	(0.061)	(0.066)	(0.083)	(0.087)
Voluntary	0.073	0.073	0.078	0.057	-0.065	0.070	0.149***	0.044	0.108
	(0.045)	(0.045)	(0.093)	(0.071)	(0.101)	(0.065)	(0.076)	(0.076)	(0.083)
Control Mean	-0.091	0.009	-0.002	-0.017	0.004	0.021	0.014	0.025	0.002
Shrp. q-val Work		0.073	0.193	0.377	0.349	0.024	0.003	0.371	0.207
Test Work=No work	0.794	0.018	0.199	0.029	0.890	0.036	0.232	0.343	0.772
Test Work=Unpaid	0.319	0.046	0.292	0.713	0.605	0.103	0.112	0.133	0.849
Observations	1188	1188	1188	1188	1188	1188	1188	1188	1188

Figures

Figure 1: Experimental Design


```
                              Blocks
                          In Refugee Camps

   Small work          No work            Volunteer              Work
   N = 40 blocks       N = 80 blocks      N=80 blocks            N = 80 blocks

  Male    Female    Male    Female    Male    Female    Male    Female
  178 HH  182 HH    359 HH  361 HH    359 HH  361 HH    365 HH  355 HH
```
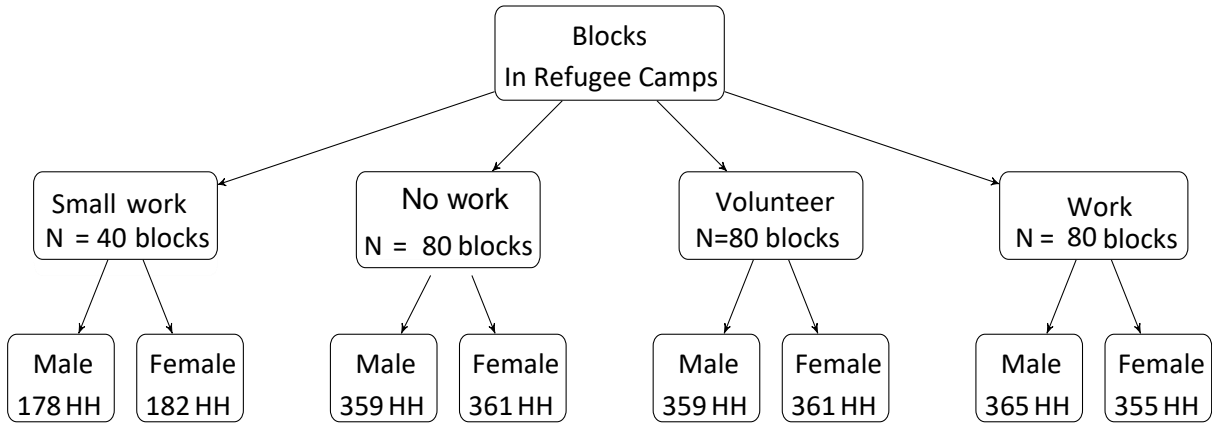
Figure 2: Pre-filled calendar

Figure 3: Work-Task

(a) Female

Name:_____ Household ID:_____ Time: _____
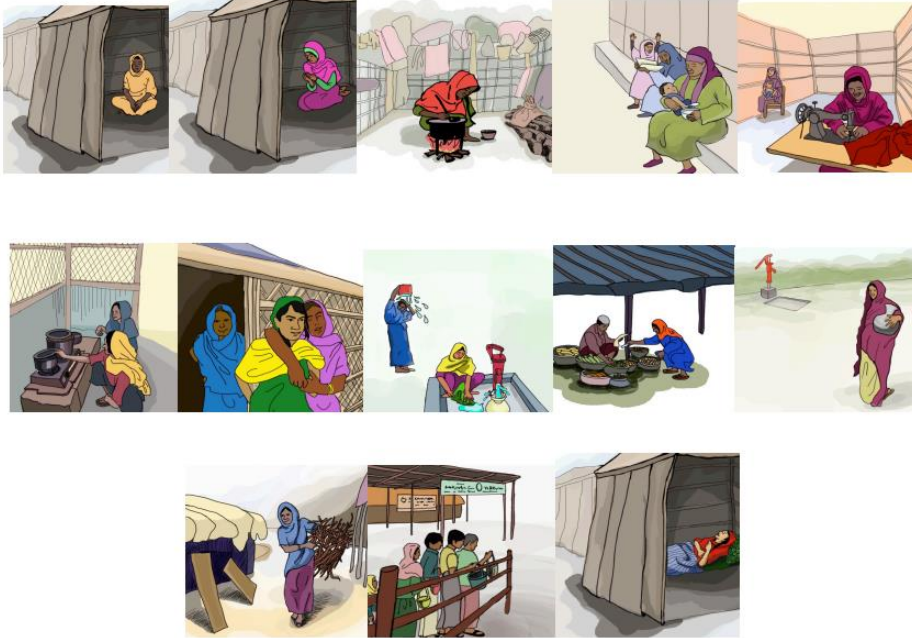
(b) Male

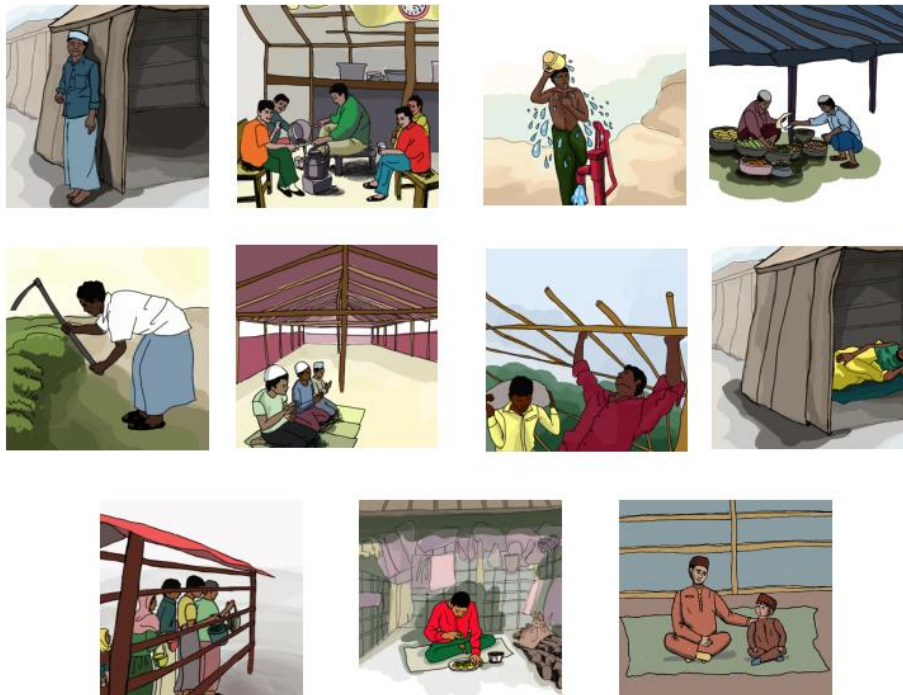Name:_____ Household ID:_____ Time: _____

Figure 4: Participation Certificate

Figure 5: Task completion