Pre-Analysis Plan: The Psychosocial Impacts of Force Idleness

Reshmaan Hussam, Erin M. Kelley, Gregory Lane, Fatima Zahra

June 10, 2020

Contents

1 Introduction

Social scientists have long posited that employment may deliver social and psychological utility beyond the value of income alone (Morse and Weiss, 1955; Jahoda, 1981; Kessler, Turner, and House, 1988). Identifying the psychosocial benefits of employment has implications for a vast range of policies, from assistance schemes for the unemployed, to government responses to forcibly displaced communities, to a future of automation and the resulting shift away from traditional forms of work. While cross-sectional evidence around this question exists (Paul and Moser, 2009; McKee-Ryan et al., 2005; Wehrle et al., 2018), this literature suffers from two key limitations. First, the challenge of selection, whereby those who are unemployed differ from the employed in ways that are likely correlated with their psychological wellbeing. Second, the inability to disentangle the mechanisms that drive the relationship between employment and wellbeing, whereby the pecuniary channel of easing resource constraints is conflated with the psychological channel of alleviating loneliness, lack of purpose, meaning, agency, or the like.

This paper presents the first real-world causal estimates of the psychosocial benefits of employment. We address both limitations in the literature by exogenously imposing employment opportunities and including a comparable group that benefits solely from the pecuniary dimension of employment. We run a field experiment in which we randomize 745 individuals of working age into three arms: (1) a control arm, in which no work is offered but a nominal fee for weekly survey participation is provided; (2) a cash arm, in which no work is offered but a large fee (equivalent to three days of daily labor) for weekly survey participation is provided; and (3) a work for cash arm, in which individuals are offered employment for an average of three days a week and are paid (in an amount equivalent to the cash arm) during the weekly survey. The work and parallel cash provisions last for eight weeks, well beyond the duration of the average daily labor opportunities that arise in our setting. Our experiment spans the peak [informal] employment season of the winter harvest.

We work with the recently displaced Rohingya refugees who reside in the largest refugee camp in the world upon the southern tip of Bangladesh. The setting is one in which nearly all households are constrained in their labor supply choice due to scarce labor market opportunities. This is both substantively and methodologically important, as providing labor market opportunities in an environment where none exist allows us to estimate the psychosocial impacts of paid employment among individuals who otherwise occupy suboptimal allocations of labor and leisure, and are therefore likely to shift behavior due to the intervention. They experience, as we term in this study, *forced idleness*.

Between August and December 2017, approximately 750,000 Rohingya fled a genocidal

campaign in Rakhine State, Myanmar, crossing the border by foot or raft into Bangladesh. Seeking to limit integration and maximize the likelihood of repatriation to Myanmar, the government of Bangladesh has made formal work illegal for the refugees. Strict restrictions on movement limit access to informal work in nearby urban centers. Refugees are desperate for more to do: many ask for *"haather kaaj,"*- colloquially, handiwork; literally, a way to keep one's hands occupied.[1] Among our sample population of male and female refugees between the ages of 18 and 45 years, individuals report spending an average of three waking hours entirely idle (the strictest measure of unoccupied time, not counting diversionary but unproductive activities).[2] Only 10% of our sample report having worked for one or more days in the previous month.

Unemployment is not only pervasive but psychologically costly: "leisure" appears to be a source of great disutility. Both intensive margin idleness and extensive margin unemployment are significantly correlated with our measure of depression (PHQ-9) at baseline: a three hour increase in hours idle is associated with a 0.12 standard deviation increase in the depression scale. Similarly, individuals who report having been unemployed the entirety of the previous month score 0.34 standard deviations higher in the depression scale, translating to a 10% higher likelihood of being moderately or severely depressed (Table A1). Naturally, these are simply baseline correlations, but suggest an important role for employment in psychological wellbeing and underscore the necessity of exogenous variation in employment in order to identify a causal relationship.

We enrich our experiment by examining the mediating roles of two features common to many forcibly displaced persons, both of which were powerfully palpable during our own time within the Rohingya camps: past violence and future uncertainty. Just as the ubiquity of forced idleness is plain upon entering the camps, likewise are refugees' memories of a traumatic past and fear of an uncertain future.

The unanticipated and indiscriminate nature of the 2017 Rohingya genocide in Myanmar presents a unique opportunity to examine the impact of past violence on the psychosocial costs of forced idleness and consequent benefits of employment. The United Nations' 2018 Human Rights Council Report by the Independent International Fact-Finding Mission on Myanmar articulates how "the operations were designed to instill immediate terror, with people woken by intense rapid weapons fire, explosions, or the shouts and screams of vil-

[1]Such expressions of the need for time to be occupied are not unique to Rohingya refugees. Syrian migrants in the Turkish Killis camp in 2017, regarded as one of the best materially-equipped refugee camps in the world, echoe these sentiments: *"We wake up, we sleep, we wake up, we sleep, we eat food... There is no purpose in a life like this. One day is like another." (McClelland, 2014)*

[2]If these are included, the average respondent in our sample spends eight hours of their waking day engaged in unproductive and/or leisure time.

lagers. Structures were set ablaze and Tatmadaw soldiers fired their guns indiscriminately into houses and fields, and at villagers." Described by one eyewitness, *"The first round of shooting was like a rain of bullets. The second round was slow as the soldiers killed the men individually. They aimed a gun at each man and shot."* The report further details how soldiers then killed those who had survived gunshot wounds with long knives, including children.

Stories such as these are not difficult to find within the Rohingya camps, and the HRC's documentation of indiscriminate violence is consistent with our baseline data: conditional on township of origin, we find that refugees who report having experienced the death of at least one family member or community member in the military raids are no different on a set of key sociodemographic observables from those who did not experience a death. 88% of the population within our sample report having experienced at least one death, with a conditional mean of 9.3 deaths. These experiences are significantly correlated with present psychosocial wellbeing: individuals who experienced a death are 0.35 standard deviations higher on the depression scale, translating to a statistically significant 25% higher likelihood of being moderately or severely depressed (Table A2, Columns 1 and 2). Our experiment exploits this quasi-random variation in past violence to investigate whether employment may serve to alleviate the psychosocial costs of past trauma. Baseline data lend support to this mediating role of employment. Unemployment is associated with a 0.6 standard deviation higher depression score, or a statistically significant 33 percentage point (87%) higher likelihood of being moderately or severely depressed, among those who experience a death relative to those who do not (Table A2, Columns 3 and 4). As above, these correlations are vulnerable to selection into employment, reinforcing the need for a field experiment to estimate the causal impact of past trauma on the psychosocial value of present employment.

The second feature of the forcibly displaced that we embed into our experimental design is the deep uncertainty that the migrant faces about his or her future. When asked what most occupies the mind during idle time, 46% of our sample volunteer concerns about the future. *"I think about where I will be in the future all the time. We are travelers here. We are anxious to go back to our own land." "Sometimes it seems that we might be transferred anywhere. But I know nothing about where we will go. I imagine sometimes that we will be shifted to another place, or thrown away in the river."*

While alleviating long-term and existential uncertainty around refugees' future is beyond the capacity of this study, we estimate the role of alleviating short-term uncertainty through the nature of the work we provide.[3] In particular, among those refugees randomized into

[3]We are motivated here by psychology literature around the value of setting short-term goals to combat depression and achieve longer term stability (Johnston et al., 2007; Crane et al., 2010; Ahrens, 1987)

the work for cash treatment, a randomized subset receive a calendar marked with every date of employment for the full eight week duration of employment. The remainder receive no such calendar and are instead informed once a week about their work schedule for the following week. We inform both groups at the outset, however, about the nature of the work, the total labor demand (24 days over eight weeks), and the total wage. By so doing, we ensure a careful mechanism experiment in which we vary only the degree of certainty with which refugees may envision their daily activities for the coming two months. We stress that in a context where tomorrow's food supplies are uncertain, daily predictability around employment over two months is likely to be a meaningful source of certainty.

Baseline data again offers suggestive evidence that considerations of the future play an important role in psychosocial wellbeing: 92% of those who report concerns of the future also report that idle time is "somewhat or very unpleasant." Notably, unemployment in the last thirty days is associated with a statistically significant 0.6 standard deviation increase in the depression score (46% higher likelihood of moderate or severe depression) among those who are occupied by the future relative to those who are not (Table A3). Yet again, these correlations are informative but susceptible to selection, necessitating exogenous variation in actual certainty about the future in order to determine its causal impact on wellbeing.[4]

We embed a final layer of randomization in an effort to cleanly identify the psychosocial mechanism at work and prevent conflation with the possible effects of future work expectations among employed individuals. To this end, we present a certificate of participation to a randomized half of our sample (across all treatment groups). This allows us to distinguish between the psychosocial value of the work itself relative to that of anticipation of future employment opportunities due to work experience. We discuss this sub-experiment in detail in the body of the paper, but stress that it is not the central objective of this study.[5]

While our experimental design is directly motivated by observations of the lived experiences of Rohingya refugees, our context of constrained labor choice in employment and forced idleness is far-reaching. Participants in the labor market often face limitations on their ability to freely trade off labor and leisure. The U.S., which typically enjoys a tight labor market relative to comparably developed countries, faces unemployment rates around

[4]This relationship between existential uncertainty and idle time is not unique to this refugee context; it echoes the observations of psychologist Victor Frankel in the concentration camps of Nazi Germany: *"A man who could not see the end of his "provisional existence" was not able to aim at an ultimate goal in life.... The unemployed worker, for example, is in a similar position. His existence has become provisional and in a certain sense he cannot live for a future or aim at a goal."* (p.70)

[5]In addition to the margins of variation described above, our setting further allows us to explore questions along two important dimensions of heterogeneity: respondent gender and baseline level of depression. Neither is the central objective of this study but both have important implications for policy; we discuss each in detail in the body of the paper.

4% as of 2019. Underemployment rates are double: 8% of the U.S. population experience "leisure" time from which they derive negative utility relative to working. Marginalized communities are particularly vulnerable to high unemployment and underemployment rates: Blacks in the U.S. faced un(under)-employment rates of 8% (12%) respectively in 2019; women, ethnic minorities, the formerly incarcerated, migrants, and the physically or mentally disabled likewise face significant restrictions in labor market opportunities globally ILO (2019). Indeed, in the current global era of COVID-19, unemployment rates within the U.S. alone have reached upwards of fifteen percent as of April 2020 and considerably higher levels for minority subpopulations (Dey et al., 2020).

This study makes four primary contributions. First, the study provides the first causal estimates in a real-world setting of the psychosocial impacts of long-term employment conditional on income, a measure that has implications well beyond those upon the refugees we study. There exists a long history of sociological work exploring the costs of long-term unemployment beyond that of income alone (Morse and Weiss, 1955; Terkel, 1974; Turner, 1995; Colic-Peisker and Walker, 2003; Jahoda, Lazarsfeld, and Zeisel, 1971; Wehrle et al., 2018). Our experiment is motivated by this observational work as well as a limited stock of empirical evidence around the psychosocial costs of idle time. We build upon the work of Bhanot, Han, and Jang (2018), who estimate the value of occupied time in a ten day lab-in-the-field experiment in Nairobi, Kenya, in which individuals are randomized into either waiting for one hour for a voucher or sorting lentils for one hour and receiving a voucher of equal value. The authors find that the latter treatment indeed improves psychological wellbeing. These results are consistent with a cross-sectional examination of workfare versus unemployment benefit recipients in Germany (Knabe, Schöb, and Weimann, 2017), in which the former reported greater wellbeing and life satisfaction despite equivalent income, as well as with the lab experiments of Hsee, Yang, and Wang (2010) and Wilson et al. (2014), both of which find that individuals are willing to pay in order to avoid being idle. Bhanot, Han, and Jang (2018) serves as valuable groundwork, as the intervention examined is of short duration and a type of work that is distant from more realistic forms of employment. We design this study as a field experiment with a plausible and long-duration form of gainful employment, which capitalizes not only on occupying idle time but also on sociability, being engaged throughout the day, and having at least nominal purpose behind the work - elements common to most, even tedious, forms of employment (Terkel, 1972).

Second, we contribute to a large policy literature around the future of work and the merits of Universal Basic Income (UBI). Widespread unemployment has implications not only for the material but also the psychosocial wellbeing of the un- and under-employed. While unemployment insurance and UBI directly address the loss of income, they do not

address the psychosocial costs that may accompany the absence of work. These costs have been explored in the sociology literature, first articulated in Jahoda, Lazarsfeld, and Zeisel (1971)'s seminal work around Marienthal, a small town in Austria that was devastated by deindustrialization in the wake of the global depression of the 1930s. As described by one woman who lost her job, *"If I could get back to the factory it would be the happiest day of my life. It's not only for the money; stuck here alone between one's own four walls, one isn't really alive."*(Jahoda, Lazarsfeld, and Zeisel, 1971). More recently, individuals who are incarcerated - as of 2019, 2.3 million within the United States alone - describe similar experiences. *"It is the dull sameness of prison life, its idleness and boredom, that grinds me down ... boredom, time-slowing boredom, interrupted by occasional bursts of fear and anger, is the governing reality of life in prison."*(Council, 2014).

Third, the study engages seriously with the underlying channels of the relationship between employment and psychosocial wellbeing. We examine through survey measures how the role of agency, sociability, and purpose may transform employment into increased wellbeing. We also estimate the role of past trauma - through quasi-random variation in exposure to violence - and future uncertainty - through exogenous variation in work schedules - in mediating the impacts we observe. While employment and job training programs are common policy levers considered for migrants and those who lack economic stability, this is the first study to both probe the mechanisms and causally estimate the mediating role of features that embody the experience of the 70 million forcibly displaced individuals in the world today (UNHCR, 2018).

Similarly, empirical work around the relationship between employment and uncertainty, conditional on income and labor demand, is scarce. There is a small literature that investigates the negative relationship between uncertainty and positive affect (Anderson et al. 2019, Carleton 2016), as well is individuals' positive willingness to pay to alleviate uncertainty (Lovallo and Kahneman, 2000). However, studies that engage the role of employment are limited. Mas and Pallais (2017) find workers are willing to pay 20% of their wages to prevent an employer from setting an irregular schedule. The setting of this study, however, is substantively different, being online workers in a developed country context where employment opportunities are not scarce. We tackle this question in a setting where unemployment and idle time are ubiquitous, the unpredictability of scheduling is unlikely to crowd-out other welfare-enhancing activities, and the uncertainty around both short-term and long-term wellbeing is psychologically crippling.

Finally, this study contributes to a small but growing literature that engages refugee populations and the forcibly displaced in field experiments (see IPA (2020) for a list of interventions). Among this set of studies, the vast majority are psychosocial support in-

terventions and the remainder explore material interventions (cash transfers, skills training, food provision, etc.). Our research is the first to examine the non-pecuniary mechanisms through which a material intervention (gainful employment) may improve psychosocial well-being. This is a valuable exercise, as aid organizations and policymakers must balance the severe material and psychological constraints that their populations face.

We note that operating in the largest refugee camp in the world, and one which formed in a matter of weeks out of a genocidal campaign of significant proportions, is not easy. Such a setting pairs the inevitably complicated politics of rebuilding society in makeshift camps, with the operations of innumerable decentralized NGOs. These complex processes engage a population that bears the trauma of their recent origins and the existential uncertainty of their future, rendering it essential for field operations consider potential unintended consequences of their activities. Neither diminishes the necessity of engaging with this population in service of welfare-enhancing policy: forcibly displaced migrants are at a historical high and projected to grow substantially in the coming decades, with the International Organization for Migration projecting up to one billion climate migrants alone by 2050 (IOM, 2014).

The remainder of the paper is structured as follows: Section 2 further describes the research context in which we operate; Section 3 outlines the experimental design; Section 4 describes our data collection processes; Section 5 proposes our main set of hypotheses and analysis plan; Section 6 reviews power calculations, and Section 7 concludes.

2 Research Context

2.1 The Rohingya

The Rohingya are an ethnic group that lives predominately in Rakhine State along the western coast of Myanmar (also known as Burma) (Blakemore, 2019). The community traces their origins back to the 15th century, when thousands of Muslims settled in the former Arakan Kingdom, which was conquered by the Burmese Empire in 1784 (Albert and Maizland, 2020). Along with the subsequent colonization of Burma by Britain in the early 1800s, came the first recorded census of Burma. The British administrative state introduced a system of ethnic classification defining 135 sub-races but did not include the Rohingya (Mahmood et al., 2017). British rule was also accompanied by a relaxation of borders within its colonial territory, encouraging migration between Burma and India (of which Bangladesh was at the time a part) in order to address labor shortages. The resulting rapid influx and economic rise of ethnic Muslims from the Bengal region of India heightened existing tensions between minority Muslims and the predominantly Bamar (Buddhist) population

within Burma (Hussam, 2019).

Following independence from Britain in 1948 and a brief period of democratic rule, General Ne Win led a successful military coup in 1962 and placed the country under military rule for the next several decades (Hussam, 2019). The new government inherited the system of ethnic classification developed from the British census and built its administrative state upon this framework. In 1982, the Citizenship Act required national identity cards specifying ethnic membership in one of the 135 recognized national races — thereby excluding, by construction, the Rohingya from citizenship (Wade, 2017). Multiple waves of suppression of the Rohingya minority ensued over the next 50 years, even as the country transitioned into democratic rule. The first major campaign of ethnic cleansing against the Rohingya occurred in 1978 when the Burmese military, tasked with performing a census of the border regions to determine citizenship, conducted indiscriminate attacks across Rohingya villages in Rakhine state. This lead to an estimated quarter million people fleeing into neighboring Bangladesh. Subsequent ethnic cleansing campaigns in 1992 and 2012 sent additional waves of Rohingya into Bangladesh (Watch, 1996).

2.2 Recent Events

On August 25, 2017, the Rohingya insurgent group Arakan Rohingya Salvation Army (ARSA) launched coordinated attacks on a military base and security force outposts across northern Rakhine, killing twelve security personnel (Hussam, 2019). Within hours, Myanmar security forces responded. As narrated by survivors in the United Nations' Human Rights Council (HRC) Report of the Independent International Fact-Finding Mission on Myanmar, abduction, detention, and gang rape in military compounds was systematic. Satellite imagery documented the destruction of at least 392 villages (40 percent of all settlements in northern Rakhine), with 80 percent burned within the first three weeks of the "clearance operations."

By October 2018, over 750,000 Rohingya refugees found themselves in a veritable city of makeshift tents along the southern tip of Bangladesh, stretching from Teknaf to Cox's Bazaar. They joined another 250,000 to 300,000 "Old Rohingya" who had left Myanmar in earlier years of ethnic cleansing. The largest and most densely populated refugee camp on earth was constructed in a matter of weeks (Hussam, 2019).

Operations within the camp are coordinated and overseen by the Bangladesh Government's Ministry of Disaster Management and Relief (MoDMR), which is represented across camps by the Refugee Relief and Repatriation Commissioner (RRRC) and within each refugee settlement by Camp-in-Charge (CiC) officials. International institutions (BRAC, UNHCR, IOM among others) actively work with the government to facilitate service de-

livery (including food, shelter, clean water, and sanitation). There are currently 34 camps in Bangladesh, each subdivided into blocks ranging in population density from 60 to 130 households. Each block is represented by a local leader (a *majhi*) who is responsible for organizing distribution efforts and serving as a liaison between humanitarian organizations, the army, the CiC, and the refugee community.

Though Bangladesh maintained open borders for the steady inflow of refugees, negotiations between the governments of Bangladesh and Myanmar around repatriation began promptly after the initial influx. Protests and international pressures forced the Bangladeshi government to delay plans for repatriation until November 2018, then, amidst further protests, indefinitely. Not wishing to encourage the long-term stay of the Rohingya, the Government of Bangladesh has enacted measures to discourage integration of refugees with host communities. In particular, refugees are not allowed to work (Bhatia et al., 2018). Many are left idle in the camp, leaving some vulnerable to various forms of human or drug trafficking (Rights Watch, 2019). Some men seek occasional employment in the informal sector outside of the camps, but this comes with significant risk as military checkpoints around the camps are abundant. An informal economy has also developed within the camps over time, providing some households with opportunities to setup small shops – though this represents a very small minority of all working-age refugees within the camps.

3 Experimental Design

3.1 Sample

The research team obtained permission from the RRRC to work in three camps in Bangladesh (5, 8W, 17), which were selected given the healthy relationship cultivated between our research partner, Innovations for Poverty Action (IPA), and the CiC officials in each location.[6] The CiC organized meetings with the local *majhis* to explain how the research team would be interacting with households in their respective blocks. Within each camp, we selected non-adjacent blocks to reduce the risk of spillovers. Finally, within each block, we selected 5 households to be part of our sample. Enumerators worked in teams of five, traveling together to two blocks per day to execute their surveying tasks. Upon entering a given block, the field team knocked on doors at random within each block, asked if the household member (pre-assigned as male or female head of household) was interested in participating in a study, and confirmed that the respondent met seven pre-established criteria: they had not worked

[6]We obtained permission to work in a fourth camp, which would have brought our sample to 1000 refugees. However, this camp posed numerous logistical burdens, which made data collection a significant challenge and led to our decision to exclude this camp from the analysis.

in the last 14 days; they were within the ages of 18-45 years; they were able and willing to work for two months inside the block; they were not the *majhi* or a member of the *majhi's* household; and they did not receive remittances from abroad. Upon meeting all criteria, the subject was enrolled into the study sample. In total, we assembled a sample of 745 individuals across the three camp sites.

3.2 Primary Intervention

We randomly assigned 149 blocks, each with five refugees enlisted in our sample, to one of two treatment arms and a control group. Figure 1 presents the experimental setup. In each case, we informed participants that the study would last eight weeks and that the field team would be checking in weekly to conduct brief five-minute surveys and provide compensation. We assigned 33 blocks to the control group (the "Small Cash" group), where participants received 50 taka (0.60 USD) per week as compensation for answering our weekly surveys. An additional 33 blocks were assigned to the "Large Cash" group, where participants received 450 taka (5.30 USD) per week as compensation for the surveys. Finally, 83 blocks were assigned to a "Cash-for-work" group, where we offered participants gainful employment. We compensated participants in this treatment arm with 150 taka (1.77 USD) per day of work. Households were assigned two, three, or four days of work per week, averaging out to 450 taka per week as in our "Large Cash" group.

In order to generate exogenous variation in short-term uncertainty, we randomly assigned individuals within work blocks to a "certain" or "uncertain" work schedule. We provided participants in the "certain" group with a pre-filled calendar that highlighted the days they were supposed to work and the days the enumerators would return to collect their work submissions and deliver payments (Figure 2). Participants in the "uncertain" group received the same work task but no calendar with pre-filled dates. They were instead informed each week about which days they would be hired for the following week. As this randomization was executed at the household level, we varied the schedule that the two groups received within a block in order to prevent uncertain participants from assuming their schedule was identical to their [potentially] certain neighbors. For example, in a particular week, the uncertain group might be assigned two days of work (Tuesday and Thursday) while the certain group could be working four days (Monday through Thursday). These schedule assignments would be reversed for another block in the same week. Across our total working sample, therefore, the schedules across weeks were (on average) identical between certain and uncertain employees.

Our employment task was designed to be easily completed by women or men of any

literacy level and working age within the study population. It was further designed to occupy the employee throughout the course of the day and in a manner that required some nominal level of engagement with his or her surrounding community. Specifically, employees were asked to engage in a data collection exercise in which they filled out a time-use sheet, reporting on the activities of twenty same-sex neighbors four times per day (centered around prayer times, which were a common source of structure for most refugees), for an average of three days per week. The neighbors that each employee selected were not identified to the researchers, ensuring that no participant felt like they were infringing on the privacy of others. This employment task was motivated by the idea (as described to our participants) that NGOs wished to understand how best to help the refugee community, and in order to do so, required an accurate picture of how refugees spend their time in the camps. This information could also be used to inform the focus of future interventions.

In order to ensure that literacy was not an impediment to completing the work, we contracted an artist to create a worksheet depicting common daily activities in the camps (sleeping, eating, lounging at a tea stall, sitting idle). We piloted the sheets extensively to ensure that all major activities were included (please see Figure 3 for a visual of the time-use sheet and activities). Upon being randomly assigned to the employment intervention, enumerators spent twenty minutes explaining the work task to households and then showed the participant a five minute video designed by the research team articulating the same; this ensured standardized comprehension across participants. Any questions that participants had regarding the task were answered at this time (and in subsequent collection survey days as needed).

We asked that households complete the work tasks on the specific days they were assigned. To ensure compliance with the work schedule, we stationed a tamper-proof box in a pre-chosen household within each block and informed participants that they should submit their tasks into the box at the end of each assigned workday. Supervisors decided which household in our sample would host the collection box (henceforth referred to as the "facilitator" household). They selected the household whose dwelling was most centrally located and who had enough space to accommodate the weekly collection surveys. These facilitators were compensated with an additional Tk. 50 for their services.[7]. At the end of each day, the facilitator would drop a piece of cardboard paper into the box, thereby separating out each day's work. The facilitator had no access to the materials inside the box.

Participants were instructed to come to the facilitator's home on the designated collection day each week. The facilitator made their home available for a few hours on this day so the enumerator could complete the check-ins with the block's five respondents and pay

[7]In our robustness checks we include an indicator for being one of these facilitators

the participants their respective amounts. In the case of work blocks, the enumerators first checked the respondents' work (the number of pages they submitted – with each page representing one of the four times per day the activity should have been completed – whether they were submitted on the correct dates, and the number of mistakes per sheet). At the end of the interaction, enumerators were instructed to look at the respondents' performance over the previous three weeks. If the work had not been completed correctly three weeks in a row, the enumerator did not pay the participant for that week. Payment occurred at the end of the interaction, once the enumerator had administered the standard weekly collection survey.

We anticipate high rates of completion among our work-task participants, based on a pilot we ran four months prior to the scale-up that confirmed refugees' eagerness to work regardless of the wage offered. However, it is important that we check non-compliance among the cash-for-work treatment arm, as failure to complete the work will result in lower wages and consequently different pecuniary benefits between the Large Cash group and the Cash for Work group. We will employ our weekly and endline surveys to check that 1) the total disbursal of cash by the end of the experiment was not significantly different between the Large Cash and Cash for Work groups; 2) that work recipients were indeed employed at a higher rate than cash recipients (i.e large cash recipients did not go out and find work for themselves); and 3) we do not see any differential attrition.

After the first four weeks of the intervention, we provided half of the blocks in our sample (across all treatments) with a certificate confirming they had engaged with the NGO for data collection activities. This treatment arm was designed to disentangle the effect of working in the *present* from the expectation of future work opportunities, where future work opportunities may be maximized with the provision of the certificate for employed participants (though we also distributed this certificate to a randomized subset in both cash groups in order to partial out any effects of reciprocity on psychosocial wellbeing). This sub-experiment offers some sense of the extent to which individuals may derive psychosocial benefits from an anticipation of future work (much like individuals who choose to engage in unpaid internships for the promise of future employment), which we wish to distinguish from psychosocial benefits from the experience work itself.

At endline, we also surprised Cash for Work households with the opportunity to engage in one additional week of work.[8] This additional feature of the experiment permits an estimation of individuals' reservation wages (which could not be elicited for the given work task until they had experience engaging in the work, and is therefore conducted at endline).

[8]We justified this surprise by informing participants that we have a some limited funds remaining for this work; this ameliorates further expectations of work after the additional week.

Moreover, by randomizing the wage offered to the employee, we can estimate their labor supply curve. In an attempt to bound the effect of reciprocity in this exercise (for example, participants may feel obligated to work for us because we had engaged them thus far), we varied how this additional week of work was framed. Specifically, among a randomized subset of participants, we emphasized that "we can find others to do the work if you don't want to" in order to stress that the participant was under no pressure to work (we had no need for his or her). In order to elicit reservation wages, we first asked households whether they would be willing to volunteer for the additional week of work. For those who said yes, we then asked them wether they would prefer to receive cash for answering a few questions instead of volunteering with us. For those who said no, we elicited their willingness to accept a given wage, using the titration method, for working an additional week. We then randomly assigned wages, allowing us to estimate the labor supply curve in our sample. Before assigning the wage, however, we also asked respondents whether they would prefer the same amount in cash for answering a few questions rather than the additional week of work.

While not the focus of the current paper, we also offered our Small Cash and Large Cash groups the opportunity to provide a week of employment at endline. This was designed to further understand how the *nature* of work affects labor supply. We designed a simple jewelry-making task which allowed us to conveniently randomize features of the work task (busyness, agency, sociability, and purposefulness) to examine the channels through which employment may improve psychosocial wellbeing. Further details of this lab-in-the-field experiment are provided in the Appendix.

4 Data Collection and Survey Instruments

4.1 Timeline and Survey Instruments

Prior to the rollout of the full experiment, the research team spent twelve months engaging in an extensive piloting of our survey instruments as well as a pilot experiment involving 300 households. Sociopolitical, emotional, cultural, and administrative complexities necessitated a deep and iterative process of developing our survey instruments. We started with standardized modules but adjusted heavily to accommodate these contextual demands, adapting or eliminating various questions from such modules which were insensitive or culturally incoherent given the experiences of the Rohingya. Major adaptations and source modules are described below, and further detail can be provided upon request. Surveys were translated and back-translated from English to Bengali to Rohingya.

Upon launching the full experiment, data were collected via a baseline, midline, and endline survey, as well as nine weekly 'collection' surveys to track a smaller number of outcomes regularly. The baseline was conducted with households prior to revealing treatment status, which was immediately followed by the midline survey to capture the short-run impacts of the treatment. Weekly surveys were conducted before participants collected their payment each week. The endline survey was conducted two days after the end of the work and cash provision period. Finally, we conducted a six-week followup survey to monitor mental health for program participants once the interventions had been concluded.

4.2 Main Outcome Variables

We define seven primary outcomes of interest: time-use, mental health, stability, physical health, cognitive function, economic decision making, and willingness to work.

4.2.1 Time-Use/Idleness

Our primary measure of time-use is the self-reported average number of hours that respondents spend idle. At baseline, we find that individuals spend an average of three hours per day sitting entirely idle (this excludes unproductive or diversionary activities like sleeping during the day or sitting at a tea stall).

While we do not prespecify the following outcomes, we seek to further understand how the interventions shift time allocation by 1) analyzing what activities respondents report substituting idle time for (enumerated as household chores, looking for a job, childcare, sitting at tea stalls, talking to friends, visiting friends, praying, or taking a nap); and 2) reviewing their detailed time-use allocation reports for the day prior to endline. We provisionally categorize activities into a hierarchy of substitutability: productive activities which are more difficult to substitute away from (bathing, market, chores, ration, eating, child-rearing), and unproductive activities which can be more easily replaced (tea-stalls, prayer, sleeping, visiting friends/relatives, playing games, playing sport, sitting idle).

4.2.2 Psychological Wellbeing

Overall Psychological Wellbeing Index We construct an overall psychological wellbeing index by aggregating the standardized outcomes listed below using inverse covariance weighting.

Depression We measure depression using the nine-question depression scale of the Patient Health Questionnaire (PHQ-9), a standardized screening tool that assesses mental and

emotional health disorders.

Stress We measure stress using an index comprised of three questions that we adapted to the refugee context from Cohen's Perceived Stress Scale (the most widely used tool for measuring the perception of stress).

Life Satisfaction Life satisfaction is measured with an adapted version of Diener's Satisfaction With Life Scale. We include four of the five statements along a six-point Likert scale.

Sociability We inquire about the interactions that participants have had throughout the day prior to the survey day. In particular, we record how many different people the respondent had a conversation with and how many of these interactions left them feeling positive.

Self-Worth We develop our own questions around self-worth rather than employing the more standard Rosenberg Self-Esteem Scale, which we found inappropriate given the Rohingya's recent experiences. Specifically, we construct an index of self-worth from three questions designed to illicit respondents' beliefs about how they contribute to their family and community. The first question invites respondents to consider the person in their community who contributes the most to their respective family, and asks the respondent where they would rank themselves relative to that individual. The second question asks respondents to rank themselves relative to the person they believe contributes the most to their community. The final question asks respondents to rate how much they are able to help their family in ways they wish to. All three questions are measured on a scale from 1-10.

Agency We capture respondents' agency in two ways. First, we build a locus-of-control index, drawn loosely from Rotter's 13 question Locus of Control Instrument. The index is comprised of four questions about the degree to which people believe that they, as opposed to external forces, have control over the outcomes in their lives. Second, we use a simple revealed preference game in which participants are offered an incentivized opportunity to either make a resource allocation decision for their community themselves or defer to another individual (an NGO worker, an "expert", or another refugee) to make the decision on their behalf.

Stability We adapt the Cantril Self-Anchoring Striving Scale (Cantril, 1965) to measure how secure respondents feel in their present lives and in the future. The questions ask respondents to consider a ladder, with the most secure life being a 10, and the least secure

life being a 0. Respondents articulate which step on the latter they feel they are in at present and where they anticipate being in five years.

4.2.3 Gender Dynamics

We construct a measure of household-level perceptions on gender and power in two ways. First, we ask about perceptions around gendered decision-making and intimate partner violence. The questions are drawn from Haushofer and Shapiro (2016), which are themselves adapted from the Demographic Health Surveys. In addition, we measure attitudes towards women's ability to work and freedom of movement by asking respondents whether they feel that women should be allowed to work, and whether this holds if the woman must work outside their respective camp block.

4.2.4 Cognitive Ability

We measure cognitive ability in two ways. First, we employ a digit-span memory test using both forward and backward sequences of numbers. Second, we ask a series of basic arithmetic problems including multiplication and division questions. Our measure of cognitive ability is the standardized index of these two measures.

4.2.5 Physical Health

We inquire about the respondent's physical health, namely whether and for how many days they have fallen sick in the thirty prior days. With 75% of the sample reporting that they were sick in the last 30 days at baseline, we focus on a measure of whether households were severely ill in the past week. Most simple infections will resolve in a week, and chronic conditions should also stabilize within a week with proper care. Our primary measure of physical health is therefore an indicator for whether respondents report being sick for seven or more days.

4.2.6 Financial Wellbeing

We ask respondents how much they have borrowed and how much they have saved in the prior 90 days. We employ this outcome measure only for comparisons of the small cash group to the large cash (and/or employment) group, as this is the margin along which there exists variation in pecuniary benefits.

4.2.7 Economic Decisionmaking

We explore economic decisionmaking along two dimensions: time and risk preferences. Time discount factors are estimated by adapting Andreoni and Sprenger (2012) convex time budget (CTB) method following Giné et al. (2018). Risk preferences (risk aversion) are measured with multiple price list decision tasks (adapted from Holt and Laury (2002) methodologies), adjusted for the Rohingya context for comprehension.

4.2.8 Willingness to Work

We capture recipients' willingness to engage in a work task at endline. We apply the incentivized Becker-DeGroot-Marschak (BDM) method among work group respondents (who now have experience with the work task), and ask them if they are willing to complete an additional week of work at various titrated wages. Pairing these responses with the number of days the respondent actually completes of additional work, we can plot the labor supply curve.

5 Hypothesis and Analysis

Across all specifications, we use double-selection LASSO to select controls for precision and we control for baseline measures of our outcomes when they are available through an ANCOVA specification. Unless otherwise specified, each specification will be run for the full set of outcomes described in Section 4.

5.1 Main Effects

5.1.1 Specification 1: What is the impact of gainful employment relative to cash alone?

$$Y_{ibc} = \beta_0 + \beta_1 LargeCash_{ibc} + \beta_2 Work_{ibc} + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

Where Y_{ibc} represents the host of outcomes outlined in Section 4 for individual i in block b and camp number c, X_{ibc} is a vector of baseline covariates, and ϵ_{ibc} is an error term clustered at the block level. We include fixed effects for camp γ_c and enumerator δ_e.[9]

We seek to understand the impact of employment ($Work$ relative to $LargeCash$) on the wellbeing of participants in our study. In order to interpret the magnitude of these effects, we benchmark them against the impact of a cash intervention ($LargeCash$ relative

[9]We follow (Di Maio and Fiala, 2019) and include enumerator fixed effects to account for the fact that respondents' answers may be influenced by the way enumerators ask more sensitive questions.

to $SmallCash$). If cash generates positive psychosocial benefits, we expect to find that $\beta_1 \geq 0$. If employment delivers psychosocial benefits equal to or beyond the benefits of income alone, we expect to find that $\beta_2 \geq \beta_1$. The ratio between these two coefficients (β_2/β_1) yields the degree to which the psychosocial benefits of employment outweigh those of cash alone.

We additionally collect weekly data on a subset of outcomes in order to observe temporal dynamics (see Table 5). The time it takes for various effects to materialize can inform the underlying mechanisms at work, and policy design around meaningful durations of employment. As such, we run the following specification:

$$Y_{ibct} = \beta_0 + \sum_{t=1}^{8} \beta_t LargeCash_{ibc} * \eta_t + \sum_{t=1}^{8} \gamma_t Work_{ibc} * \eta_t + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{itbc}$$

Where Y_{ibct} represents the stress index, sociability, cognitive ability, and physical health, η_t represents a dummy for the weekly visit number t, and γ_c, δ_e, X_{ibc}, ϵ_{itbc} are as defined above.

5.1.2 Specification 2: How does past violence mediate the psychosocial impacts of employment?

Do the impacts of employment vary by the degree to which an individual experiences violence in his or her recent past? We answer this question by running the following specification on the sample who received the large cash intervention or the employment intervention:

$$Y_{ibc} = \beta_0 + \beta_1 Violence_{ibc} + \beta_2 Work_{ibc} + \beta_3 Work * Violence_{ibc} + \theta_l + \gamma_c + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

We focus exclusively on measures of mental health and stability for individual i in block b and camp number c (Y_{ibc}); θ_l represents fixed effects for the respondent's township of origin in Myanmar; and γ_c, δ_e, X_{ibc}, ϵ_{ibc} are as defined above. If past violence indeed mediates the psychosocial impact of employment, we expect to find that $\beta_3 \neq 0$. Our observations from baseline data suggest that $\beta_3 \geq 0$: those who experienced greater violence in the past are likely to experience greater psychosocial benefits from an employment intervention.

We define exposure to violence, $Violence$, as an indicator for whether the household experienced at least one death in his or her family or community during the 2017 exodus from Myanmar. Two pieces of evidence suggest that exposure to violence was indeed quasi-random: our universe of sociodemographic measures appear balanced between those who did or did not experience at least one death; and reports by the Human Rights Council repeatedly articulate the indiscriminate nature of the Myanmar military's "Clearance Operations". We

consider each in turn.

Table 3 presents a balance table between those individuals who reported at least one death in their family or community and those who reported none, conditional on the respondent's township of origin. We present the results including (1) township-level fixed effects and (2) fifty square mile grid-level fixed effects. We select township fixed-effects because all refugees originate from one of three townships (Maungdaw, Buthidaung, or Rathedaung – see map in Figure 4 of Appendix) and it is not clear whether the military chose to attack one administrative boundary for political expediency or internal alliances before another. Second, in the absence of available data on military movement but presuming that the military could not cover all regions at once, we group origin villages into grid cells of fifty square miles (although balance is robust to a variety of cell sizes).[10] Across all available time-invariant sociodemographic measures we collect, features appear balanced between those who did and did not witness at least one death under both the township-level fixed effect specification and the grid-cell level fixed effect specification (Columns 3 and 4, respectively). It is worth noting that sociodemographic measures are balanced even absent controls for location of origin (Column 5). Out of precaution, however, we choose to impose township fixed effects in all violence-related regressions (which we prefer over the grid-cell fixed effects as it reduces room for arbitrary experimenter determination and relies purely on existing geographical/administrative boundaries).

One may be still concerned about strategic movement of the military if it is correlated with unobservables that also mediate the relationship between employment and psychosocial outcomes (for example, perhaps wealthier regions were attacked first, leading to those residents witnessing more violence in Myanmar and potentially benefiting more from employment, given their past social status, in Bangladesh). While there are no news sources that track the geographic movement of the military during the Clearance Operations, and available satellite images on burning villages are too infrequent to deduce military movement, a large body of detailed qualitative evidence suggests that violence was indiscriminate at local levels. The Human Rights Council of the United Nations commissioned the Independent International Fact-Finding Mission on Myanmar in September of 2018. The 441 page report interviews several hundred victims and eyewitnesses and describes in detail the nature of violence perpetrated in August through December of 2017 in Rakhine State. Among other similar excerpts, the report describes how *"many Rohingya were killed or injured by indiscriminate shooting. Rohingya villages were approached without warning, usually from more*

[10] In the grid cell specifications, 66 observations are dropped because the villages the respondents reported coming from could not be matched to existing geocoded village names. All townships could be matched to existing geocoded townships.

than one direction, and often in the early morning, by armed Tatmadaw soldiers.... The operations were designed to instill immediate terror, with people woken by intense rapid weapons fire, explosions, or the shouts and screams of villagers. Structures were set ablaze and Tatmadaw soldiers fired their guns indiscriminately into houses and fields, and at villagers." In Figure 5, we present a selection of excerpts from the report that further describe the indiscriminate nature of this violence.

5.1.3 Specification 3: How does alleviating future uncertainty affect psychosocial wellbeing?

Does reduced uncertainty about the daily prospects of the next two months improve refugees' psychosocial wellbeing? To answer this, we restrict our sample to only those individuals who received employment, and compare those who received a calendar outlining their complete work schedule ($Calendar$) to those who did not.[11] We run the following specification:

$$Y_{ibc} = \beta_0 + \beta_1 Calendar + \lambda_s + \zeta_b + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

We limit our pre-specified outcomes (Y_{ibc}) to those collected at midline (all of which were obtained immediately after the baseline survey once the respondent was randomly allocated a work schedule or not), as we anticipate that the effect of providing a schedule is highly time-sensitive and challenging to capture through longer term and more generalized measures of wellbeing. ζ_b represents block-level fixed effects (as calendar provision was randomized at the individual, not block, level), and λ_s represents fixed effects for the work schedule each employee (knowingly or unknowingly) received. δ_e, X_{ibc}, and ϵ_{ibc} are as defined above. If provision of a calendar for the two months of employment indeed has an impact on psychosocial and financial decision-making measures, we expect to find that $\beta_1 \neq 0$. If increased certainty is mediating this effect, we anticipate $\beta_1 \geq 0$ for the stability and financial outlook outcomes. However, the effects on risk and time preferences remain directionally ambiguous.

5.2 Secondary Analyses

5.2.1 Gender

Our exploration of differential effects by gender is motivated by literature in sociology around the loss of work and gender identity (Payne (1998); Schrijvers (1997)), some of which suggests

[11] Note that both groups were explicitly informed about the total number of days they would work, the nature of the work, and the total wage received. This allows us to isolate the pure impact of uncertainty around daily activities on psychosocial wellbeing.

that job loss leads to greater male aggression in the home due to a greater sense of powerlessness and lack of agency (Annan and Brier (2010); Heltberg, Hossain, and Anna Turk (2012); Kabeer (2015); Ondeko and Purdin (2004); Wirtz et al. (2014); Patinkin (2014)). This work is consistent with recent evidence that the COVID-19 lockdowns of 2020, which increased the presence of males in the home due to work-from-home regulations and job loss, was correlated with increased incidence of domestic and intimate partner violence (Economist (2020); Godbole (2020)). We are further motivated by literature in economics around how employment may raise the household bargaining power of females (a more thorough review of which can be found in McKelway (2020)).

We focus on four outcomes: power perceptions (and the subcomponents of perceptions around intimate partner violence and broader gender norms),[12] time use, mental health outcomes, and stability. In addition to the prespecified primary measure of time-use (on self-reported hours idle), we are also interested in exploring (though we do not prespecify) men and women's time-use allocation between productive and unproductive activities. We suspect that women may be forced to substitute away from productive activities when employed, while men may find themselves substituting away from idle time. We also examine mental health outcomes, where the total effect on female psychosocial wellbeing is ambiguous. On the one hand, women may experience smaller benefits to mental health if they must forgo other productive activities in order to perform their work tasks. Conversely, women who are able to leave the home may be less vulnerable to domestic abuse within the household and potentially gain a greater sense of agency, translating to a boost in mental health and sense of stability. Drawing from the sociology literature on gender identity, we anticipate the impacts of employment on the mental health and sense of stability of males to be unambiguously positive. We run the following regression:

$$Y_{ibc} = \beta_0 + \beta_1 LargeCash_{ibc} + \beta_2 LargeCash * Male_{ibc} +$$

$$\beta_3 Work_{ibc} + \beta_4 Work * Male_{ibc} + \zeta_b + \delta_e + X_{ibc} + \varepsilon_{ib}$$

Where Y_{ibc} represents power perceptions, time-use, mental health, and stability for individual i in block b and camp number c; and where ζ_b, δ_e, X_{ibc}, ϵ_{ibc} are as defined above. This regression allows us to examine the effects of a cash transfer (β_2) and gainful employment (β_4) on our outcomes of interest. A comparison of the two coefficients demonstrates whether the non-pecuniary effects of employment are differentially greater for males relative

[12]We prespecify the power perception index as the primary outcome because it is a standardized index drawn from the DHS, but we also suspect that an index as broad as this will be unlikely to shift from our intervention, and therefore are especially interested in how its subcomponents around intimate partner violence may respond to the intervention.

to females.

5.2.2 Levels of depression

Our investigation of the heterogeneous effects of employment by baseline depression levels is motivated by psychological literature that explores the potential vicious cycle of depression, in which those who are especially depressed lack the ability to recall positive pasts (Teasdale, 1983) or conceive of possible futures (Roepke and Seligman, 2016), thereby sinking further into depressed states.

This idea is echoed in recent work by Haushofer and de Quidt (2019), who describe how many depressed patients "are frequently unable to derive pleasure from otherwise enjoyable activities, suggesting a change in preferences." However, we argue that it is *ex ante* ambiguous whether the most depressed individuals in our sample will react the most or the least to the employment intervention: while those who are severely depressed may not have the psychological foundations necessary to benefit from the potential psychosocial gains that employment can offer, those who are already in a psychologically healthy space may be least likely to need and therefore benefit from the non-pecuniary benefits of employment. As such, we explore how baseline levels of depression mediate the magnitude of the effect of employment on psychosocial wellbeing, using the medical definitions of mild, moderate, moderately severe, and severe depression as derived from the PHQ-9 scale. We pool the latter two groups for purposes of power (as 9% of our sample qualifies as severely depressed at baseline) and run the following regression:

$$Y_{ibc} = \beta_0 + \beta_1 LargeCash_{ibc} + \sum_{d=1}^{4} \beta_d LargeCash_{ibc} * DepressionLevel_d +$$

$$\beta_3 Work_{ibc} + \sum_{d=1}^{4} \gamma_t Work_{ibc} * DepressionLevel_d + \zeta_b + \delta_e + X_{ibc} + Deaths_{ibc} + \varepsilon_{ib}$$

Where Y_{ibc} represents mental health and stability outcomes for individual i in block b and camp number c, $DepressionLevel_d$ is a dummy for each level of depression (mild, moderate, and moderately severe and severe pooled; omitted category is not depressed), where ζ_b, δ_e, X_{ibc}, ϵ_{ibc} are as defined above, and $Deaths_{ibc}$ is a control for the total number of deaths that occurred among the respondent's friends and family members.

We make no claims about the nature or source of measured depression, though we recognize there is certainly a wide variation in the degree of violence witnessed and perpetrated upon the self among this population. The above is a purely exploratory analysis, but an important one when considering the policy implications of employment interventions within

a context such as ours.

5.2.3 Future income streams

One challenge we anticipate with the interpretation of our results is that current employment may make future employment more likely, and therefore carry monetary benefits beyond those of the immediate income received (either through the relationship formed with the NGO, or through a boost in the beneficiary's 'resume' which may make them more appealing to other potential employers). We view this as highly unlikely (both in actuality and in expectation) within our context given the limited employment opportunities as well as our repeated reminders (at baseline, two weeks before endline, and endline) that the work opportunity we were providing would last only eight weeks, and we would not be conducting any additional activities in the camp thereafter.

Recognizing that this potential future pecuniary benefit of employment is impossible to rule out, however, we randomized the provision of paper certificates which provided documentation of the beneficiaries' involvement with our project: an explicit boost to their resume. These certificates were signed by our enumerators and included the following text: *"Certificate: This acknowledges that I engaged with Pulse Bangladesh to do data collection."* (Figure 6). In order to control for any reciprocity effects, we also provided identical certificates to a randomized subset of the cash-only arms.[13] If employed individuals derive psychosocial benefits from an expectation of future work, the certificate should make this expectation maximally salient. This should provide some sense of how concerned one may be about a conflation of purely psychosocial mechanisms with [future] pecuniary mechanisms.

The effect of the embedded certificate randomization is estimated via the following regression:

$$Y_{ibc} = \beta_0 + \beta_1 LargeCash_{ibc} + \beta_2 Work_{ibc} + \beta_3 LargeCash * Certificate_{ibc} +$$

$$\beta_4 Work * Certificate_{ibc} + \beta_5 Certificate_{ibc} + \zeta_b + \delta_e + X_{ibc} + \varepsilon_{ibc}$$

Where Y_{ibc} represents our mental health index and the stability outcome for individual i in block b and camp number c, and where ζ_b, δ_e, X_{ibc}, ϵ_{ibc} are as defined above. We are interested in whether $\beta_4 \geq \beta_3 \geq 0$: 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_3).

[13] The wording of the certificate was made such that it could be applied to both arms; cash-only arms participated in weekly surveys along with all other experiment participants, so technically also engaged in data collection for our project.

5.3 Multiple Hypothesis Testing

Our programs can affect a wide range of psycho-social and economic outcomes. As a result, we will account for multiple hypotheses by computing False Discovery Rate (FDR) q-values. The primary groups of outcomes are listed above, and for each of these outcomes we will construct indices (where possible using inverse covariance weighting) and report the p-values and sharpened q-values.

6 Statistical Power

We present power calculations for a subset of the primary outcomes described above: four measures comprising our psychosocial wellbeing index (PHQ-9, Life Satisfaction, Stress, and the Stability Index), one measure of cognitive ability (digit span), and one measure of physical health (severe health problems). All power calculations are based on a hypothesis test with a 5% significance level comparing a sample of 165 individuals within 33 clusters of five people each (the "Large Cash" group) to a sample of 415 individuals within 83 clusters of five people each (the "Work for Cash" group). Outcome means and standard deviations are calculated using our baseline data. Each table presents power calculations for variously-sized treatment effects and across a range of inter-cluster correlations (ICC). We vary the ICC from 0.01 to 0.2 as baseline ICCs were very low (ranging from 0.000 to 0.06).[14] Overall, we are well powered to detect treatment effects of approximately 10%-15% (0.2 - 0.25 standard deviations). These numbers are comparable to those found in the literature for similar indices and outcomes: Haushofer and Shapiro (2016) find a 0.26 standard deviation increase in their primary wellbeing index as a result of treatment.

Mental Health Outcomes: Respondents' PHQ-9 at baseline has a mean of 8.3 and a standard deviation of 4.4, however, for interpretation purposes we will examine standardized effects on PHQ. We present power calculations for treatment effects ranging from 0.15 to 0.3 standard deviations. Our calculations suggest we are well powered to detect a treatment effect between 0.25 to 0.3 standard deviations (12.5% to 15%). For the life satisfaction and stress indices, we are powered to detect between a 10% and 12.5% treatment effect off of a mean of 11.09 – 4.8 s.d. and 9.58 – 3.8 s.d., respectively.

[14] This may not be surprising due to the fact that the blocks are newly formed communities where individuals were quasi-randomly located after fleeing Myanmar.

Power Calculation for PHQ

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
0.15 Std Dev	0.38	0.34	0.30	0.24
0.20 Std Dev	0.60	0.54	0.48	0.39
0.25 Std Dev	0.79	0.73	0.66	0.56
0.30 Std Dev	0.91	0.87	0.82	0.71

Power Calculation for Life Satisfication

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
10%	0.69	0.63	0.57	0.47
12.5%	0.87	0.82	0.76	0.65
15%	0.96	0.93	0.89	0.80
17.5%	0.99	0.98	0.96	0.91
20%	1.00	1.00	0.99	0.96

Power Calculation for Stress Index

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
7.5%	0.53	0.48	0.42	0.34
10%	0.78	0.72	0.65	0.54
12.5%	0.93	0.89	0.83	0.73
15%	0.98	0.97	0.94	0.87
17.5%	1.00	0.99	0.98	0.95

Stability: The Stability Index has a mean of 14.3 and standard deviation of 4.5 in baseline data. This index has the highest ICC of our main outcomes (0.06). Despite this relatively high ICC, we are powered to detect a 10% treatment effect.

Power Calculation for Stability Index - Work Treatment

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
5%	0.41	0.37	0.32	0.26
7.5%	0.74	0.68	0.62	0.51
10%	0.94	0.90	0.85	0.76
12.5%	0.99	0.98	0.96	0.91
15%	1.00	1.00	0.99	0.98

The Stability Index is also a core outcome of interest for the certain vs. uncertain work schedule treatment arms. For this experiment, we compare the 208 respondents in the certain arm to the 207 respondents in the uncertain arm. Randomization of the schedule was done at the individual level within blocks, removing concerns about clustering and the ICC. The table below indicates that we are also powered to detect a 10% increase in the stability index for this comparison.

Power Calculation for Stability Index - Certainty Treatment

	Effect Size				
	5%	7.5%	10%	12.5%	15%
Power	0.32	0.60	0.84	0.96	0.99

Cognitive Ability: We present power calculations for the sum of the digit span tests (one forwards, one backwards). The average person was able to remember a total of 6.07 digits with a standard deviation of 1.48. We present power calculations for a treatment effect ranging from a change of 0.2 to 0.6 digits. We are well powered to detect a treatment effect of at least 0.4 digits in this outcome.

Power Calculation for Digit Span

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
0.2 Digits	0.35	0.31	0.27	0.22
0.3 Digits	0.65	0.59	0.52	0.43
0.4 Digits	0.88	0.83	0.77	0.66
0.5 Digits	0.97	0.95	0.92	0.84
0.6 Digits	1.00	0.99	0.98	0.95

Physical Health: Our measure of physical health is the probability of being severely ill, or having a health problem last more than seven days in the past month. At baseline, 29.3% of respondents reported being severely ill. We present power calculations for treatment effect sizes ranging from 5 to 15 percentage points. Our calculations suggest that we are powered to detect an effect size of 12.5 percentage points.

Power Calculation for Persistent Health Problem

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
5% points	0.22	0.20	0.18	0.15
7.5% points	0.43	0.38	0.33	0.27
10% points	0.66	0.60	0.53	0.44
12.5% points	0.84	0.79	0.72	0.61
15% points	0.94	0.91	0.87	0.77

Financial Wellbeing: We have baseline data for two financial outcomes of interest: current savings and current debt from loans. For these outcomes, we are interested in comparing the Small Cash group to the Large Cash group. As before, all power calculations are based on a hypothesis test with a 5% significance level. However, we are now comparing a sample of 165 individuals (the "Large Cash" group) in 33 blocks to a sample of 165 individuals within 33 blocks (the "Small Cash" group). The average respondent only has 200 taka of savings (s.d. 600) at baseline, and 2,267 taka of debt (s.d. 3151). The ICC for financial outcomes is quite low: 0.000 for savings and 0.012 for loans. However, for robustness we continue to

show power calculations with ICC's ranging from 0.01 to 0.2. The tables below show that we are powered to detect changes in savings levels of approximately 100%. Given the very low levels of current savings, and the relatively sizeable cash transfers we provide in the Large Cash intervention, we find this to be a reasonable effect size. For loans, we are powered to detect a 40% change in debt levels, which is also plausible given our intervention.

Power Calculation for Savings

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
30%	0.16	0.14	0.13	0.11
40%	0.25	0.22	0.19	0.16
50%	0.36	0.32	0.28	0.23
75%	0.66	0.60	0.54	0.44
100%	0.89	0.84	0.78	0.68

Power Calculation for Loans

	Inter-Cluster Correlation			
Effect Size	**0.01**	**0.05**	**0.1**	**0.2**
30%	0.59	0.53	0.47	0.39
40%	0.83	0.78	0.71	0.60
50%	0.96	0.93	0.88	0.79
75%	1.00	1.00	1.00	0.99
100%	1.00	1.00	1.00	1.00

References

Ahrens, Anthony H. 1987. "Theories of Depression: The Role of Goals and the Self-Evaluation Process." *Cognitive Therapy and Research* 11 (6):665–680.

Albert, Eleanor and Lindsay Maizland. 2020. "What Forces Are Fueling Myanmar's Rohingya Crisis?" https://www.cfr.org/backgrounder/rohingya-crisis.

Andreoni, James and Charles Sprenger. 2012. "Estimating Time Preferences from Convex Budgets." *American Economic Review* 102 (7):3333–3356.

Annan, Jeannie and Moriah Brier. 2010. "The Risk of Return: Intimate Partner Violence in Northern Uganda's Armed Conflict." *Social Science & Medicine (1982)* 70 (1):152–159.

Bhanot, Syon P., Jiyoung Han, and Chaning Jang. 2018. "Workfare, Wellbeing and Consumption: Evidence from a Field Experiment with Kenya's Urban Poor." *Journal of Economic Behavior & Organization* 149:372–388.

Bhatia, Abhishek, Ayesha Mahmud, Arlan Fuller, Rebecca Shin, Azad Rahman, Tanvir Shatil, Mahmuda Sultana, K. A. M Morshed, Jennifer Leaning, and Satchit Balsari. 2018. "The Rohingya in Cox's Bazar." *Health and Human Rights* 20 (2):105–122.

Blakemore, Erin. 2019. "Who Are the Rohingya People?" https://www.nationalgeographic.com/.

Cantril, Hadley. 1965. *The Pattern of Human Concerns*. Rutgers University Press.

Colic-Peisker, Val and Iain Walker. 2003. "Human Capital, Acculturation and Social Identity: Bosnian Refugees in Australia." *Journal of Community & Applied Social Psychology* 13 (5):337–360.

Council, National Research. 2014. *The Growth of Incarceration in the United States*. 10.17226/18613. Washington, DC: The National Academies Press.

Crane, Catherine, Thorsten Barnhofer, Emily Hargus, Myanthi Amarasinghe, and Rosie Winder. 2010. "The Relationship between Dispositional Mindfulness and Conditional Goal Setting in Depressed Patients." *British Journal of Clinical Psychology* 49 (3):281–290.

Dey, Matthew, Mark Loewenstein, Anne Polivka, and David Piccone. 2020. "Demographics, Earnings, and Family Characteristics of Workers in Sectors Initially Affected by COVID-19 Shutdowns : Monthly Labor Review: U.S. Bureau of Labor Statistics." Tech. rep.

Di Maio, Michele and Nathan Fiala. 2019. "Be Wary of Those Who Ask: A Randomized Experiment on the Size and Determinants of the Enumerator Effect." *The World Bank Economic Review* .

Economist, The. 2020. "Domestic Violence Has Increased during Coronavirus Lockdowns." *The Economist* .

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang. 2018. "Revising Commitments: Field Evidence on the Adjustment of Prior Choices." *The Economic Journal* 128 (608):159–188.

Godbole, Tanika. 2020. "Domestic Violence Rises amid Coronavirus Lockdowns in Asia." *DW* .

Haushofer, Johannes and Jonathan de Quidt. 2019. "Depression Through the Lens of Economics: A Research Agenda." In *The Economics of Asset Accumulation and Poverty Traps*, The Economics of Poverty Traps. The University of Chicago.

Haushofer, Johannes and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics* 131 (4):1973–2042.

Heltberg, Rasmus, Naomi Reva Hossain, and Carolyn Anna Turk. 2012. *Anatomy of Coping: Evidence from People Living through the Crises of 2008-11*. Policy Research Working Papers. The World Bank.

Holt, Charles A. and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5):1644–1655.

Hsee, Christopher K., Adelle X. Yang, and Liangyan Wang. 2010. "Idleness Aversion and the Need for Justifiable Busyness." *Psychological Science* 21 (7):926–930.

Hussam, Reshmaan. 2019. "The Rohingya Refugee: Past, Genocide, Future."

ILO. 2019. "World Employment and Social Outlook – Trends 2019." Tech. rep., International Labor Organisation.

IOM. 2014. "IOM Outlook on Migration, Environment and Climate Change." Tech. rep., International Organization for Migration.

IPA. 2020. "IPA Compliation: Experimental Studies on the Forcibly Displaced." Tech. rep., Innovations For Poverty Action.

Jahoda, Marie. 1981. "Work, Employment, and Unemployment: Values, Theories, and Approaches in Social Research." *American Psychologist* 36 (2):184–191.

Jahoda, Marie, Paul Lazarsfeld, and Hans Zeisel. 1971. "Marienthal: The Sociography of an Unemployed Community." *Social Forces* 50 (4):550–551.

Johnston, Olwyn, Satinder Kumar, Kathleen Kendall, Robert Peveler, John Gabbay, and Tony Kendrick. 2007. "Qualitative Study of Depression Management in Primary Care: GP and Patient Goals, and the Value of Listening." *The British Journal of General Practice: The Journal of the Royal College of General Practitioners* 57 (544):872–879.

Kabeer, Naila. 2015. "Gender, Poverty, and Inequality: A Brief History of Feminist Contributions in the Field of International Development." *Gender & Development* 23 (2):189–205.

Kessler, Ronald C., J. Blake Turner, and James S. House. 1988. "Effects of Unemployment on Health in a Community Survey: Main, Modifying, and Mediating Effects." *Journal of Social Issues* 44 (4):69–85.

Knabe, Andreas, Ronnie Schöb, and Joachim Weimann. 2017. "The Subjective Well-Being of Workfare Participants: Insights from a Day Reconstruction Survey." *Applied Economics* 49 (13):1311–1325.

Lovallo, Dan and Daniel Kahneman. 2000. "Living with Uncertainty: Attractiveness and Resolution Timing." *Journal of Behavioral Decision Making* 13 (2):179–190.

Mahmood, Syed S., Emily Wroe, Arlan Fuller, and Jennifer Leaning. 2017. "The Rohingya People of Myanmar: Health, Human Rights, and Identity." *The Lancet* 389 (10081):1841–1850.

Mas, Alexandre and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12):3722–3759.

McClelland, Mac. 2014. "How to Build a Perfect Refugee Camp." *The New York Times* .

McKee-Ryan, Frances, Zhaoli Song, Connie R. Wanberg, and Angelo J. Kinicki. 2005. "Psychological and Physical Well-Being during Unemployment: A Meta-Analytic Study." *The Journal of Applied Psychology* 90 (1):53–76.

McKelway, Madeline. 2020. "Women's Employment in India: Intra-Household and Intra-Personal Constraints." *Working Paper* .

Morse, Nancy C. and Robert S. Weiss. 1955. "The Function and Meaning of Work and the Job." *American Sociological Review* 20 (2):191–198.

Ondeko, Roselidah and Susan Purdin. 2004. "Understanding the Causes of Gender-Based Violence — Forced Migration Review." *Forced Migration Review* .

Patinkin, J. 2014. "Refugees in Their Own Land: South Sudan Camps Breed Idleness, Frustration." .

Paul, Karsten and Klaus Moser. 2009. "Unemployment Impairs Mental Health: Meta-Analyses." *Journal of Vocational Behavior* 74:264–282.

Payne, Lina. 1998. "Food Shortages and Gender Relations in Ikafe Settlement, Uganda." *Gender & Development* 6 (1):30–36.

Rights Watch, Human. 2019. ""Are We Not Human?" — Denial of Education for Rohingya Refugee Children in Bangladesh." https://www.hrw.org/report/2019/12/03/are-we-not-human/denial-education-rohingya-refugee-children-bangladesh.

Roepke, Ann Marie and Martin E. P. Seligman. 2016. "Depression and Prospection." *The British Journal of Clinical Psychology* 55 (1):23–48.

Schrijvers, Joke. 1997. "Internal Refugees in Sri Lanka: The Interplay of Ethnicity and Gender." *The European Journal of Development Research* 9 (2):62–82.

Teasdale, John D. 1983. "Negative Thinking in Depression: Cause, Effect, or Reciprocal Relationship?" *Advances in Behaviour Research and Therapy* 5 (1):3–25.

Terkel, Studs. 1974. *Working; People Talk about What They Do All Day and How They Feel about What They Do*. New York:Pantheon Books.

Turner, J. B. 1995. "Economic Context and the Health Effects of Unemployment." *Journal of Health and Social Behavior* 36 (3):213–229.

UNHCR. 2018. "Global Trends - Forced Displacement in 2018." Tech. rep., United Nations High Commission for Refugees.

Wade, Francis. 2017. *Myanmar's Enemy Within*. London, UK: Zed Books Ltd.

Watch, Human Rights. 1996. "The Rohingya Muslims: Ending a Cycle of Exodus? —." https://www.hrw.org/report/1996/09/01/rohingya-muslims-ending-cycle-exodus.

Wehrle, Katja, Ute-Christine Klehe, Mari Kira, and Jelena Zikic. 2018. "Can I Come as I Am? Refugees' Vocational Identity Threats, Coping, and Growth." *Journal of Vocational Behavior* 105:83–101.

Wirtz, Andrea L., Kiemanh Pham, Nancy Glass, Saskia Loochkartt, Teemar Kidane, Decssy Cuspoca, Leonard S. Rubenstein, Sonal Singh, and Alexander Vu. 2014. "Gender-Based Violence in Conflict and Displacement: Qualitative Findings from Displaced Women in Colombia." *Conflict and Health* 8 (1):10.

Tables

Table 1: Balance (All)

	(1) SmallCash	(2) LargeCash	(3) Work	(4) (1) vs. (2)	(5) (1) vs. (3)	(6) (2) vs. (3)
Female	0.32	0.29	0.30	0.24	0.45	0.49
Married	0.82	0.81	0.76	0.34	0.04	0.31
Age	28.39	29.03	28.01	0.74	0.41	0.17
Household size	4.99	5.23	5.14	0.52	0.61	0.78
Formal education	0.48	0.44	0.51	0.70	0.14	0.07
Past Ag. Work	0.62	0.66	0.65	0.74	0.64	0.92
Math ability (index)	2.61	2.59	2.58	0.90	0.43	0.38
Digit Span Score (Total)	5.94	6.07	6.13	0.63	0.18	0.35
Wellbeing (index)	-0.12	0.05	0.03	0.15	0.03	0.83
Life Satisfaction	11.04	10.85	11.21	0.62	0.05	0.22
Self-worth (relative)	13.95	14.62	13.96	0.40	0.32	0.93
Worked in the last month	0.11	0.11	0.10	0.88	0.49	0.39
Worked in Myanmar	0.72	0.72	0.76	0.57	0.81	0.38
Hours Idle (avg)	2.97	3.31	3.01	0.99	0.39	0.46
Idle Feelings	1.66	1.73	1.67	0.31	0.06	0.66
Locus of Control	7.44	7.40	7.61	0.92	0.27	0.36
Power Perceptions	10.87	10.70	10.86	0.19	0.85	0.07
Work Perceptions	3.77	3.60	3.76	0.61	0.63	0.31
Persistent Illness (>7)	0.30	0.33	0.28	0.89	0.26	0.24
PHQ Scale	8.19	8.73	8.20	0.31	0.80	0.18
Sev. Depressed	0.05	0.11	0.09	0.06	0.05	0.59
Stress (index)	9.48	9.94	9.49	0.24	0.96	0.18
Number of conversations	16.13	16.35	16.48	0.85	0.68	0.46
Number of conversations +	9.25	8.96	9.94	0.34	0.69	0.07
Number of conversations -	3.45	4.04	3.84	0.45	0.40	0.88
Family Injuries (Burma)	1.79	1.70	1.68	0.58	0.26	0.72
Observations	165	165	415			

Columns (1), (2), and (3) show the average value of the variable in the respective treatment arm. Column (4) shows the p-value of the difference in means between the Small Cash and Large Cash treatment groups. Column (5) shows the p-value of the difference between the Small Cash and Work treatments, while column (6) shows the p-value between Large Cash and Work.

Table 2: Balance (Certainty)

	(1) Uncertain	(2) Certain	(3) Uncertain vs. Certain
Female	0.32	0.28	0.65
Married	0.75	0.77	0.88
Age	27.64	28.38	0.47
Household size	5.06	5.22	0.86
Formal education	0.50	0.52	0.18
Past Ag. Work	0.69	0.62	0.52
Math ability (index)	2.57	2.59	0.75
Digit Span Score (Total)	6.13	6.14	0.72
Wellbeing (index)	0.12	-0.06	0.43
Life Satisfaction	11.45	10.98	0.86
Self-worth (relative)	14.00	13.91	0.47
Worked in the last month	0.10	0.11	0.91
Worked in Myanmar	0.73	0.78	0.83
Hours Idle (avg)	3.00	3.01	0.90
Idle Feelings	1.68	1.65	0.62
Locus of Control	7.20	8.03	0.09
Power Perceptions	10.91	10.81	0.11
Work Perceptions	3.79	3.73	0.72
Persistent Illness (>7)	0.25	0.31	0.15
PHQ Scale	8.14	8.26	0.86
Mod. Depressed	0.37	0.36	0.48
Stress (index)	9.16	9.82	0.49
Number of conversations	16.23	16.73	0.64
Number of conversations +	9.83	10.05	0.70
Family Injuries (Burma)	1.69	1.66	0.98
Observations	207	208	

Columns (1) and (2) show the average value of the variable in the respective treatment arm. Column (3) shows the p-value of the difference in means between the Uncertain and Certain treatment groups.

Table 3: Balance (Violence)

	(1) No Violence	(2) Violence	(3) No Vio. vs. Vio.	(4) No Vio. vs. Vio., Town FE	(5) No Vio. vs. Vio, Grid FE
Married	0.82	0.78	0.70	0.68	0.71
Age	27.80	28.54	0.22	0.25	0.19
Household size	5.16	5.17	0.65	0.81	0.60
Formal education	0.42	0.49	0.29	0.21	0.29
Math ability (index)	2.66	2.57	0.13	0.10	0.13
Past Ag. Work	0.58	0.66	0.15	0.12	0.13
Observations	60	430			

Columns (1) and (2) show the average value of the variable for respondents who did and did not have a family member killed in Myanmar. All difference in means test control for gender because violence was targeted differently between men and women. Column (3) shows the p-value of the difference in means with no additional controls. Column (4) reports p-values while controlling for township fixed effects, while column (5) includes fixed effects using 55 by 55 kilometer grid cells for respondent location of origin in Myanmar.

Table 4: Outcome Variable Descriptions

Psychological Well-being	
PHQ9	The standardized total score of 9 questions from the Patient Health Questionnaire-9 (PHQ9)
Life Satisfaction Index	A standardized average of survey responses to four questions from Diener's standardized scale, responses made along a seven-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]?"
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.
Self Worth Index	The standardized total score from the responses on a scale from 1 to 10 to three questions: "Think of a person you know who you [respect / think helps] the most in your [family / community]. If that person is a 10 where would you put yourself?"
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..."
Allocation Decision Game	Indicator (yes / no) for response to an offer to participate an allocation committee to decide how money is spent. Participants are offered the opportunity to make a resource allocation decision for their community or have another individual (an NGO worker, an "expert", or another refugee) make the decision.
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]"
Physiological Index	A standardized inverse-covariance weighted average of the above indices.
Gender Dynamics	
Gender Perceptions - Work	The standardized total score of two questions regarding women's work, "How often would you agree that women should be allowed to work for a living [inside /outside] the block?"
Gender Perceptions - Violence (IPV)	The standardized total score of five questions regarding norms for intimate partner violence (IPV) from the Demographic and Health Survey (DHS).
Financial Wellbeing	
Savings	Response to the question "How much money do you currently have in savings?" During the collection surveys (midlines) this question instead asked "How much money did you save in the past week?"

Borrowing	Total amount of money the household has borrowed.
Economic Decision Making	
Risk Preference	Measured using incentivized responses to the multiple price list decisions adapted from Holt-Laury and Sprenger (2002).
TIme Preference	Measured by adapting Andreoni and Sprenger's (2011) convex time budget method following Giné et al. (2018).
Other Outcomes	
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. Only the arithmetic problems were included in midline.
Physical Health	An indicator for prolonged health problems that persisted for more than one week over the past month. Coded from a question asking respondents "In the past 30 days, how many days were you sick?". For the collection surveys (midline), this question was modified to ask "How many of the last 7 days did you feel sick?"

Table 5: Outcome Variable Collection Periods

	Basline	**Midline**	**Weekly**	**Endline**
Psychological Well-being				
PHQ9	X			X
Life Satisfaction Index	X			X
Stress Index	X		X	X
Sociability (Total)	X		X	X
Sociability (Positive)	X		X	X
Self Worth Index	X			X
Locus of Control	X			X
Allocation Decision Game		X		X
Stability Index		X		X
Physiological Wellbeing Index	X			X
Gender Dynamics				
Gender Perceptions - Work	X			X
Gender Perceptions - Violence (IPV)	X			X
Financial Wellbeing				
Savings	X		X*	X
Borrowing	X			X
Economic Decision Making				
Risk Preference		X		X
Time Preference		X		X
Other Outcomes				
Cognitive Ability	X		X*	X
Physical Health	X		X*	X

*Physical Health, Savings, and Cognitive Ability are measured differently at midline than at baseline or endline.

Figures

Figure 1: Experimental Design

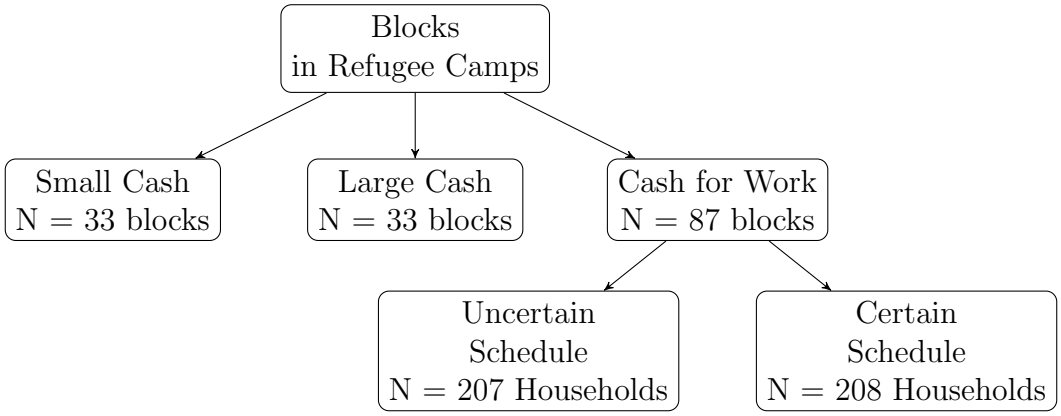


Figure 2: Pre-filled calendar

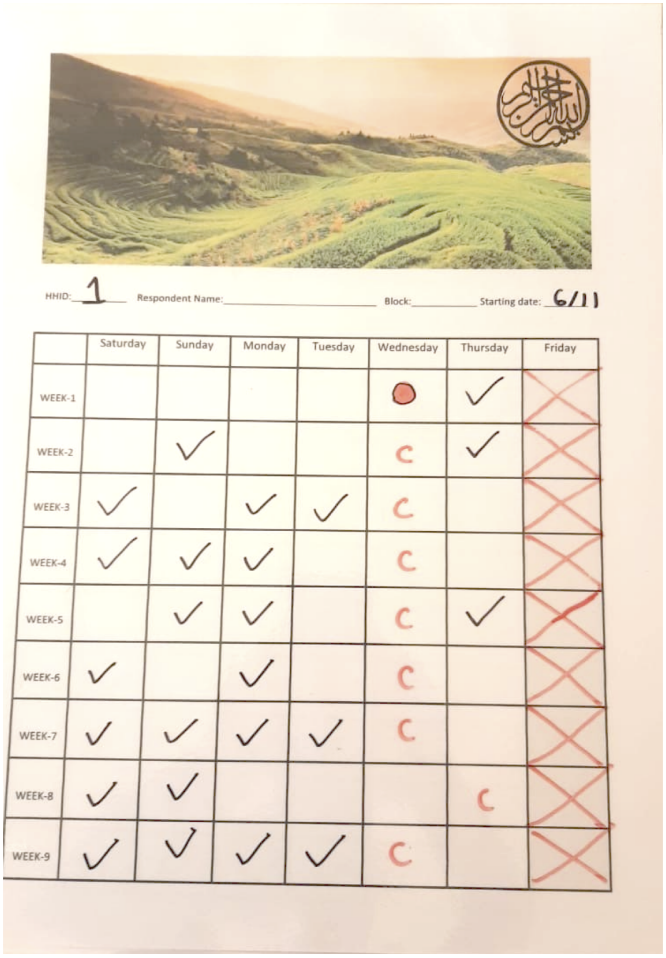
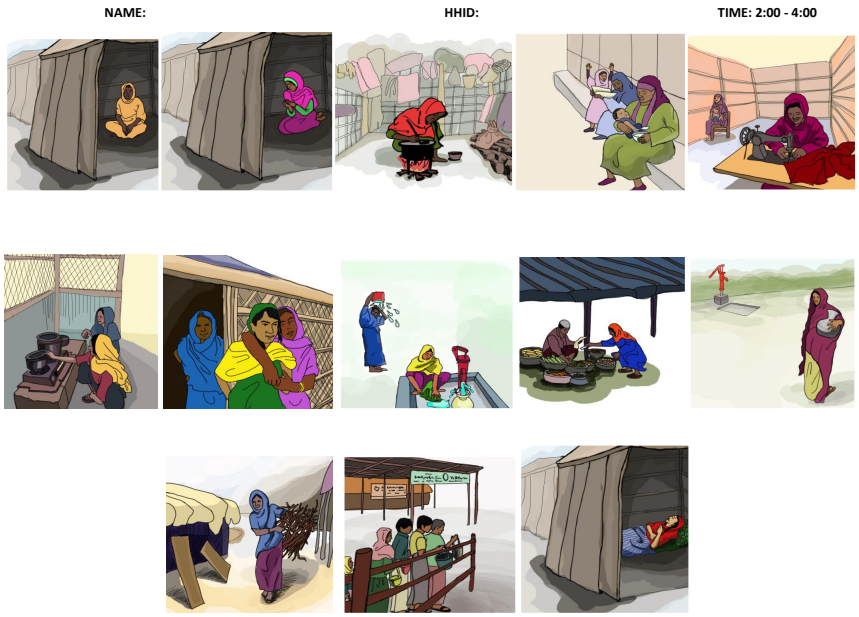


Figure 3: Work-Tasks

(a) Female



(b) Male

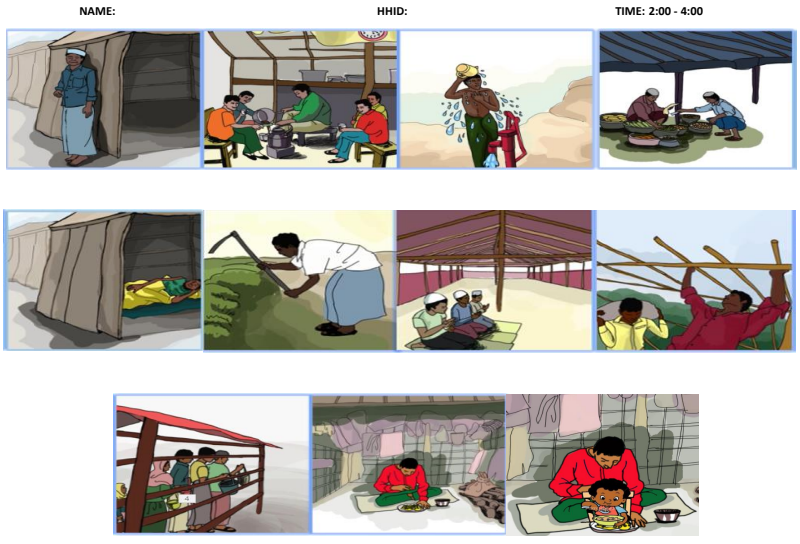


Figure 4: Map of Participant Origins

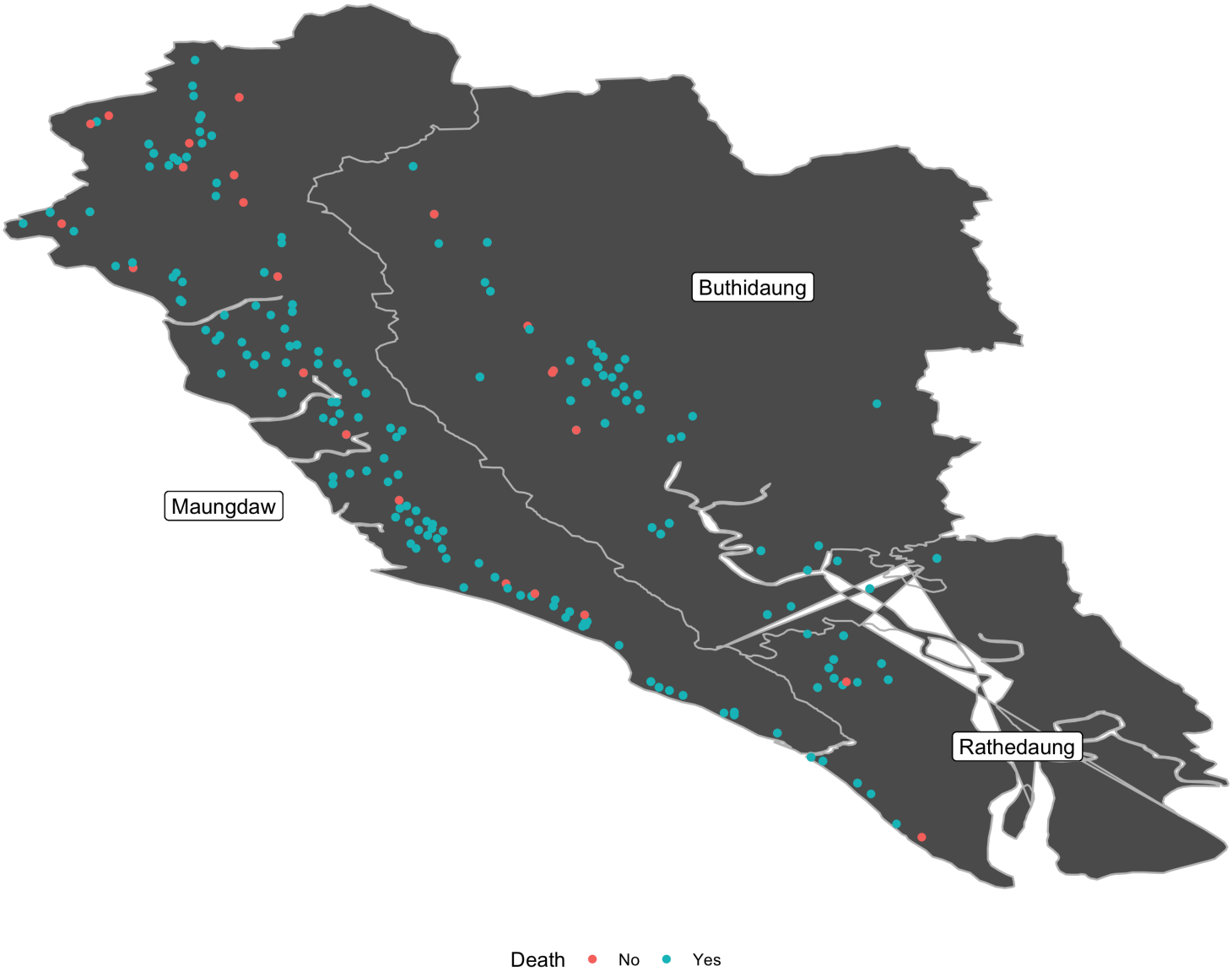


Figure 5: Excerpts from Human Rights Council Report

The following is a compilation of excerpts drawn from the United Nations' Human Rights Council Report on Myanmar regarding the "Clearance Operations" in Rakhine State executed by the Myanmar military (referred to below as the *Tatmadaw*) in late August and early September of 2017. These excerpts describe the indiscriminate nature of the violence perpetrated against the Rohingya during these operations. We caution the reader as several of these excerpts are difficult to read. We have left out the most graphic descriptions but direct the reader to the report itself (A/HRC/39/CRP.2) for further evidence of the random nature of violence during the Operations.

- During subsequent operations in villages and towns, the Tatmadaw did also not attempt to distinguish civilians from military objectives. Such indiscriminate attacks resulted in civilian men, women and children being injured or killed, with large numbers of civilians being driven away from their homes and villages. (P.35)

- Information therefore strongly indicates that airstrikes and shelling were used indiscriminately as a more general tactic in the context of "clearance operations," in essence attacking the civilian population as a whole as opposed to being used against specifically identified military targets. (P.35)

- The operations were designed to instill immediate terror, with people woken by intense rapid weapons fire, explosions, or the shouts and screams of villagers. Structures were set ablaze and Tatmadaw soldiers fired their guns indiscriminately into houses and fields, and at villagers. (P.178)

- Many Rohingya were killed or injured by indiscriminate shooting. Rohingya villages were approached without warning, usually from more than one direction, and often in the early morning, by armed Tatmadaw soldiers.... Members of the security forces, primarily Tatmadaw soldiers of the Western Command and the 33rd and 99th LIDs, shot assault rifles towards the Rohingya villages from a distance, not targeting any particular military objective or making any distinction between ARSA fighters and civilians. Men, women and children were all shot at. Many victims referred to the volume of gunfire, with some describing it as "raining bullets." Many were shot and killed or injured while attempting to flee. (P.205)

- One young girl described the operation in Maungdaw Township: "When the soldiers came to my village, we all ran, and they shot at us. We were around 50 people, and maybe half of us were shot. The people shot fell down while they were running. Some died and some escaped. Somehow, I escaped."' (P.205-206)

- One man from Kyein Chaung village tract, known in Rohingya as Boli Bazar, in northern Maungdaw Township explained the circumstances in which his daughter was killed: "I don't know how many people died that day. The military, they were just

shooting at whomever. They were shooting at people whenever they saw them, on the streets or in the houses. When they were shooting, there was no time to look back and care for those who were shot. As people were running, they were shooting at them. That is how my daughter died. She was hit fleeing. I couldn't go back and carry her." (P.206)

- Some Rohingya villagers who could not flee, or who sought shelter inside their houses, were also shot and killed or injured, when bullets penetrated thatched roofs and bamboo walls. Villagers were shot in other locations where they had found shelter, including through rapid arms fire into forested hills where they had fled. (P.206-207)

- The Mission has provided detailed accounts above of corroborated mass killings perpetrated in the villages of Min Gyi, Maung Nu, Chut Pyin, Gu Dar Pyin, the villages of Koe Tan Kauk. Dozens, and in some cases hundreds, of men, women and children were killed. Additional organized mass killings are likely to have taken place. Witnesses reported seeing bodies of large numbers of Rohingya, including those with gunshot and machete wounds, as well as decapitated heads, in burned villages en route to Bangladesh. (P.207)

- Rohingya fleeing the "clearance operations" also faced violent attacks at border crossing points, resulting in loss of life and serious injuries. Soldiers opened fire on groups of Rohingya at or close to border crossing points, including large numbers gathered on the shores of the Bay of Bengal or Naf River, while waiting to cross into Bangladesh.2005 A man from Nga Yant Chaung village tract, Buthidaung Township, described arriving at the Naf River in mid-September 2017 and being fired upon by soldiers. Some of the people ran; others, like him, lay on the ground. He said that 25 people were killed, including three of his relatives. (P.208)

- Soldiers also shot at boats carrying Rohingya to Bangladesh, resulting in further casualties. One witness explained how the boat she was in was shot at by soldiers as it crossed the Naf River, killing three men and two women. Another witness described her experience while waiting for a boat: "Soldiers started shooting, so we crawled away and lay down behind the plants in the mud. I saw many people being shot at. Dead bodies of men, women and children were floating in the river." (P.208-209)

- Another feature of the "clearance operations" was the widespread destruction of Rohingya homes and villages, causing further death and injury through burning. Houses were burned both manually using flammable liquid and matches, and by the use of "launchers," weapons firing a munition that explodes upon impact. This latter method in particular meant that victims were often caught by surprise and had little time to escape. (P.209)

- Landmines, planted by the Tatmadaw in and around Rohingya villages as part of the "clearance operations" also caused death and injury. On or around 26 August 2017, a group of Tatmadaw soldiers approached Sin Oe Pyin (Ywar Gyi) hamlet, in Maung Gyi Taung village tract, Buthidaung Township. They systematically planted mines along the main road to the village, with one villager describing them as being placed

"15 feet apart." Once the operations began, the landmines killed and injured many who tried to flee.2037 As one villager described, "The mines were put at the entrance of the village, that is the only way out so when people were running they stepped on them and died." Another recalled: "Some people were running and were killed by the mines, as they didn't know that they were planted there. Others were hit by the mines as they were coming back from the field. My 18-year old relative died from an explosion coming back from the paddy field just in front of my house." (P.211)

Figure 6: Participation Certificate

A Appendix

A.1 Tables

Table A1: Mental Health and Idleness

	(1) PHQ Scale (Std. Dev.)	(2) PHQ Scale (Std. Dev.)	(3) Mod. Depressed	(4) Mod. Depressed
Hours Idle (avg)	0.039*		0.007	
	(0.021)		(0.010)	
No work in last month		0.343***		0.103*
		(0.118)		(0.058)
Constant	-0.725***	-0.909***	0.185**	0.115
	(0.190)	(0.206)	(0.093)	(0.101)
Mean of DV	0.00	0.00	0.38	0.38
Controls	Y	Y	Y	Y
Obs	745	745	745	745

Notes: The table shows cross-section relationship between four mental health outcomes and two measures of idleness at baseline. Hours Idle is the self-reported average number of hours spent sitting idle (excluding recreational or other unproductive activities). "No work in last month" is an indictor variable for those who self-reported as having no paid work in the past month. All regressions include controls for respondent age, martial status, household size, education, being a head of household, and gender.

Table A2: Violence and Work

	(1) PHQ Scale (Std. Dev.)	(2) Mod. Depressed	(3) PHQ Scale (Std. Dev.)	(4) Mod. Depressed
Any Death	0.353***	0.092*	-0.177	-0.198
	(0.109)	(0.053)	(0.292)	(0.143)
No work in last month			-0.176	-0.180
			(0.289)	(0.142)
No Work X Death			0.603*	0.333**
			(0.314)	(0.154)
Constant	-0.881***	0.134	-0.726**	0.293*
	(0.196)	(0.096)	(0.321)	(0.157)
Mean of DV	0.00	0.38	0.00	0.38
Controls	Y	Y	Y	Y
Obs	745	745	745	745

Notes: "No work in last month" is an indictor variable for those who self-reported as having no paid work in the past month. "Any Death" is an indicator variable for the respondent having a family member or friend who was killed in Myanmar. All regressions include controls for respondent age, martial status, household size, education, being a head of household, and gender.

Table A3: Future Orientation and Work

	(1) PHQ Scale (Std. Dev.)	(2) Mod. Depressed	(3) PHQ Scale (Std. Dev.)	(4) Mod. Depressed
Think of Future	0.188**	0.062*	-0.343	-0.090
	(0.073)	(0.036)	(0.216)	(0.106)
No work in last month			0.062	0.022
			(0.160)	(0.079)
No Work X Future			0.600***	0.172
			(0.228)	(0.112)
Constant	-0.702***	0.174**	-0.770***	0.150
	(0.181)	(0.088)	(0.229)	(0.113)
Mean of DV	0.00	0.38	0.00	0.38
Controls	Y	Y	Y	Y
Obs	745	745	745	745

Notes: "No work in last month" is an indictor variable for those who self-reported as having no paid work in the past month. "Think of Future" is an indicator variable for respondents who report thinking primarily about the future during time spent idle. All regressions include controls for respondent age, martial status, household size, education, being a head of household, and gender.

A.2 Secondary Intervention

To better understand which non-pecuniary dimensions of employment are most valued by employees, we offered our small and large cash groups the opportunity to engage in a week's worth of work after the completion of our endline survey. We designed a bracelet-making task and randomized respondents, by block, into four variations of the work experience: busyness (our control group), agency, sociability, and purpose. In the first group, we informed participants that a supplier from Dhaka had ordered 10,000 beaded bracelets from us. We provided a prototype bracelet that we asked respondents to recreate (with a specific color pattern). We estimated that each bracelet would require five minutes to complete, and respondents were tasked with completing 25 bracelets per day. We asked that participants come to an assigned work space in the block to do their work, but that each person work in silence to avoid distractions and mistakes. We provided each participant with the three bead colors they would need to complete the pattern.

In the second group, we introduced an element of "sociability" to the work by requiring that participants interact with one another to create the bracelets. Each participant was only provided with one bead color and needed to exchange with others to create the required pattern. In the third group, we introduced an element of "agency" by requiring participants to develop their own patterns with all three bead colors, asking them to be as creative as they wished and design beautiful products. In the fourth group, we additionally introduced an element of "purpose" by informing participants that their bracelets would be donated to children in the camps.

Participants completed the work in the facilitator's home (the site at which they had previously collected cash payments and completed surveys). We provided participants with four bags labeled with their respondent ID and instructed them to place their completed work into one of the bags at the end of each day. Like the time-use sheets, they then dropped this bag in a tamper-proof box in the facilitator's home, and the facilitator separated each day's work with a piece of cardboard. This ensured that the employment activity was of meaningful length not only within the day but over the course of the week.

To test whether different features of the work affect participants' willingness to work, we run the following regression:

$$Y_{ibc} = \beta_0 + \beta_1 Sociability + \beta_2 Purpose + \beta_3 Agency + \zeta_b + \delta_e + X_{ibc} + \varepsilon_{ib}$$

In which Y_{ibc} represents (1) the wage at which the respondent is willing to work; (2) the number of days worked; and (3) the number of bracelets completed correctly. γ_c represents block-level fixed effects and δ_e, X_{ibc} and ϵ_{ibc} are as defined above. If provision of a calendar for

the two months of employment indeed has an impact on psychosocial and financial decision-making measures, we expect to find that $\beta_1 \neq 0$. If increased certainty is mediating this effect, we anticipate $\beta_1 \geq 0$ for the stability and financial outlook outcomes. However, the effects on risk and time preferences remain directionally ambiguous.