Forced Displacement, Mental Health, and Child Development: Evidence from the Rohingya Refugees*

Asad Islam[†] Tanvir Ahmed Mozumder[‡] Tabassum Rahman[§] Tanvir Shatil[¶] Abu Siddique[‖]

December 27, 2021

Abstract

Forced displacement is a major driver of mental disorders among refugees worldwide. Poor mental health of adult refugees, particularly mothers, is also considered a risk factor for the psychological well-being and development of their children. In this paper, we experimentally examine the extent to which a multifaceted psychoeducation program promote psychological well-being of refugee mothers and physical, socioemotional, and cognitive development of their children under the age of 2 years. Through a clustered randomized controlled trial among the severely persecuted Rohingya refugees residing in Bangladesh, roughly 3,500 mother-child dyads were given weekly psychosocial support for a year that includes psychoeducation and parenting counseling for mothers and play activities for children. The intervention was largely successful and led to: (i) improvements in the psychological trauma and depression of both mothers and children, (ii) improvements in communication, gross-motor, problem-solving, and social skills of children, and (iii) reductions in stunting, underweight, and wasting among children. A mediation analysis suggests that the mental health of both mothers and children were important channels of impact on children's anthropometric, cognitive, and non-cognitive development. The intervention also caused the mental health of children to be more aligned with the mental health of their mothers, implying policies targeting the mental well-being of forcibly displaced mothers can be an important stepping stone to developing psychological resilience among their children that can help them grow into well-rounded, healthy adults.

JEL: I15, J15, O12, O15

Keywords: Mental health, forced displacement, ethnic violence, early childhood development, randomized experiment, psychoeducation, Rohingya.

*We thank Fabian Dehos, Steven Dieterle, Eliana La Ferrara, Soledad Giardili, Patricia Justino, Tatiana Kornienko, Catherine Porter, Saurabh Singhal, and Ana Nuevo-Chiquero for many helpful comments and suggestions. We also thank seminar participants at Australian National University, Bath, Duisburg-Essen, Edinburgh, Lancaster, Monash, Technical University of Munich, and conference participants at CEPR/LEAP Workshop in Development Economics, Bavarian Micro Day Workshop, Maastricht-BEPS, Measuring Development at Berkeley, Mental Health Workshop at Max Planck Institute, MWIEDC at Northwestern, NEUDC at Boston, Workshop on Conflict, Migration, and Displacement, Young Economists Meeting at Masaryk, and Field Days Workshop at Cologne for many helpful comments. We also thank Erum Marium and Sakila Yesmin from the BRAC Institute of Educational Development and BRAC fieldworkers in Rohingya refugee camps for extensive support and collaboration for the entire project. This project has received ethics approval from BRAC University (ref no. 2019-028-ER) and is registered at the AEA RCT Registry (AEARCTR-0004516).

[†]Department of Economics and CDES, Monash, Australia; J-PAL. Email: asadul.islam@monash.edu

[‡]BRAC Institute of Governance & Development, Bangladesh. Email: tanvir.mozumder@bracu.ac.bd

[§]School of Medicine & Public Health, Newcastle, Australia. Email: tabassum.rahman@uon.edu.au

[¶]BRAC Institute of Governance & Development, Bangladesh. Email: tanvir.shatil@bracu.ac.bd

[‖]Department of Political Economy, King's College London, UK; CDES. Email: abu.b.siddique@kcl.ac.uk

1 Introduction

More than 82 million people have been forcibly displaced worldwide—a number that has doubled in 20 years (UNHCR, 2018). Forced displacement has been a major driver of mental disorders among refugees due to both pre-migration experiences, such as exposure to violence, conflict, and persecution, and migration and post-migration conditions that include separation, physical danger, unemployment, inadequate housing, financial strain, hostility, and uncertainty among others (Fazel et al., 2005; Porter & Haslam, 2005; Lindert et al., 2009; Stillman et al., 2009; Steel et al., 2009; Miller & Rasmussen, 2010; Sommer et al., 2018). The most vulnerable groups within the forcibly displaced are children and women (UNHCR, 2018). The reason being that early life crises among children can induce mental health problems during adulthood and can also impair human capital accumulation (Adhvaryu et al., 2019; Currie & Stabile, 2006; Heckman et al., 2006). Women also suffer from a complex traumatic syndrome due to trauma caused by political terror and violence (Herman, 2015). As a result of poor mental health, women that are mothers can also adversely affect their children's development and nutrition provision (Patel et al., 2004; Rahman et al., 2008; Mensah & Kiernan, 2010). Therefore, the impact of pre and post-migration experiences on the mental health of vulnerable refugees are substantial with possible detrimental economic consequences (Kessler & Frank, 1997; Currie, 2009; Ridley et al., 2020).

One severely persecuted and forcibly displaced population are the *Rohingya* people of Myanmar. They have long been disenfranchised and the 2017 incidence of genocide, community violence, and rape caused the mass displacement of about 750,000 people from Myanmar to neighboring Bangladesh (UNHCR Emergencies, 2019). Currently, one million *Rohingya* people resides in overcrowded refugee camps in southeastern Bangladesh, making them the world's largest stateless population (Bhatia et al., 2018). Adverse experiences of sexual violence and displacement from Myanmar, combined with heightened sexual abuse and intimate partner violence in refugee camps, and uncertainty and unemployment in Bangladesh, is gradually deteriorating mental health of *Rohingya* women (The Lancet, 2019). Alongside women, 80,000 *Rohingya* children are estimated to have severe mental distress and the number is also expected to grow (The Lancet, 2019). In addition, acute malnutrition, poor immunization, anaemia, and stunting are also prevalent among children (Hossain et al., 2019). Besides, poor access to health services, scarce food, and basic shelters make the *Rohingya* people vulnerable to infectious diseases (Ahmed et al., 2018), particularly during the coronavirus pandemic (Islam & Yunus, 2020), which is also harming their mental health. Mental health problems make up 13 percent of overall global disease (Collins et al., 2011), which causes an estimated economic loss of 55.5 million disability-adjusted life years in developing countries (World Health Organization, 2008). Therefore, with an increasing number of people around the world fleeing due to civil war, state failure, natural disaster, poverty, etc. (UNHCR, 2016, 2018), causing severe trauma, stress, and depression among the displaced, it is important to understand how persecution and displacement induced mental disorders can be effectively improved in low-income settings.

In this study, we investigate whether providing psychoeducation—an integration of psychotherapeutic and educational interventions to help people cope with mental illnesses by providing education, information, and a safe space for sharing (American Psychological Associa-

tion, 1995; Lukens & McFarlane, 2004)—to persecuted *Rohingya* refugee mothers can improve their psychological well-being and socioemotional, physical, and cognitive development of their children under the age of 2 years. In partnership with the BRAC Institute of Educational Development and the BRAC Institute of Governance and Development, Bangladesh, we conducted a clustered randomized controlled trial in 17 large *Rohingya* refugee camps located in southern Bangladesh. Each camp comprises many *blocks*—communities with many households within. We randomly selected 252 blocks from an universe of over 2,000 blocks and assigned the psychoeducation intervention to 137 blocks, while the remaining 114 blocks were not given any intervention (the control group). Within each block, we randomly assigned an average of fourteen mother-child pairs to the intervention. Mothers received weekly psychosocial support in groups for almost a year (44 weekly sessions) that includes psychoeducation on mental health and parenting counselling for mothers, and culturally appropriate play activities for children. The intervention ran from July 2019 to October 2020, along with a baseline and an endline survey.[1]

We find that the intervention significantly improved psychological trauma and depression of treated mothers. Specifically, mothers that received psychoeducation experienced a 0.23 standard deviation (SD) reduction in trauma and 0.14 SD reduction in depression relative to mothers in the control group that did not receive the psychoeducation intervention. Among the mothers that were identified to have psychological trauma and depression at baseline, we observe a sizeable improvement in their mental well-being—effects that are comparable to the short-run impact of a cognitive behavioral therapy on maternal depression (Baranov et al., 2020). In addition, mothers also experienced an improvement in their self-reported level of happiness (0.12 SD) and sense of belongingness in the host community (0.18 SD) following the intervention, but we do not observe any significant impact on their aspirations for the future. Furthermore, mentally unhealthy mothers that received the psychoeducation treatment caught up to, and often surpassed, the mental well-being of the mentally healthy mothers in the control group following the intervention. Thus, the intervention was largely successful at lifting treated refugee mothers out of psychological distress and potentially prevented them from developing long-term illness, such as dysthymia or persistent depressive disorder.

With respect to child outcomes, children in the treatment group experienced statistically significant improvements in terms of psychological trauma (0.10 SD) and depression (0.12 SD) relative to children in the control group. We also observe that the intervention improved their communication skills (i.e., speech and language development) by 0.23 SD, gross-motor skills (i.e., physical activities and whole-body movements) by 0.18 SD, problem-solving skills (i.e., learn to play with toys and solve puzzles) by 0.18 SD, and personal-social skills (i.e., care for themselves and interactions with others) by 0.13 SD relative to untreated children. We also estimate the impact on the prevalence of malnutrition among children, such as underweight (weigh too little or too thin for age), stunting (skeletal growth retardation), and wasting (acute undernutrition). We find that children in the treatment group experienced: (i) a large increase in weight-for-age z-score (or WAZ) by 0.64 SD (20% or 635 grams higher weight) than children in the control group, which also translates to 7 percentage points (or 10%) reduction in being underweight and 16 percentage points (or 26%) reduction in being severe underweight; and, (ii) a large increase

[1]Because of COVID-19 restrictions, the endline survey was conducted over the phone.

in height-for-age z-score (or HAZ) by 0.52 SD (19% or 1.58 centimeters taller) than children in the control group, which also translates to 7 percentage points (or 10%) reduction in stunting and 13 percentage points (or 22%) reduction in severe stunting. Both improvements in weight and growth are also reflected on the incidence of wasting among children, where treated children experienced a large increase in weight-for-height z-score (or WHZ) by 0.51 SD (21% higher) than children in the control group, which also translates to 8 percentage points (or 14%) reduction in wasting and 13 percentage points (or 30%) reduction in severe wasting. Thus, our psychoeducation intervention was largely successful at improving nutritional, socioemotional, physical, and cognitive development of the Rohingya children. These impacts should be very attractive to policymakers—particularly in low- and middle-income countries hosting refugees—as resources for refugees are often limited there and more-costly child development support approaches in refugee camps, such as food or cash transfers, pediatric support, special health care needs, etc., are often unavailable and not prioritized by host governments.

Our unique design with mother-child dyads also allows us to examine the intergenerational correlation of mental health between mothers and children, and the causal impact of the intervention on the mother-child mental health gap. We find that the mental health of mothers are strongly and positively correlated with the mental health of their children at baseline and this pairwise correlation strengthened following the intervention, implying more aligned mental health between mothers and their children at endline. In fact, the intervention was successful at reducing the mental health gap (both in terms of psychological trauma and depression) between mothers and their children. This suggests that interventions targeting the mental well-being of mothers can be an important stepping stone to developing psychological resilience among their children. This is very important in contexts where psychosocial support facilities for infants and small children are scarce and unavailable.

To understand the mechanisms at play behind children's development, we first estimate the effect of the intervention on various intermediate outcomes. We find strong evidence that the intervention increased mothers' daily involvement on their children by about 1.5 hours but muted impacts on fathers' time input, implying maternal time investment on children is a potential channel. We also find that treated mothers are less likely to allow their children to play or walk barefoot (which prevents hookworm infections and exposure to various bacterial and fungal organisms) and engage in negative parenting, suggesting improvements in mothers' health behaviors toward their children are other potential mechanisms behind their children's development. However, other maternal behaviors and practices, such as breastfeeding time, feeding frequency, seeking others' help to babysit, and discourage fathers to smoke indoors, are not likely channels of impact. We next consider a formal mediation analysis following Heckman et al. (2013); Heckman & Pinto (2015) to understand: (i) how much of mother's mental well-being and children's own development contributed to children's improvements in mental well-being; and, (ii) how much of mothers' and children's mental health improvements contributed to children's development outcomes. We find that roughly 60% of depression improvements and 80% of psychological trauma improvements among children were due to improvements in their mothers' mental health and their nutritional, socioemotional, physical, and cognitive development. We also find that 20% of improvements in child development outcomes can be explained by mothers' and children's mental health improvements. The remaining effects were

due to the direct effect of the intervention and other unobserved channels.

Finally, we examine heterogeneity in treatment effects and find the treatment effect to be stronger among mothers from households that were highly exposed to violent conflicts and persecution in Myanmar as opposed to mothers from households that had little to no exposure to violent conflicts. However, we do not find any robust evidence for heterogeneity in treatment effects by mothers' exposure to camp abuse or children's gender.

To summarize, we provide evidence of a clean, causal impact of a home-based psychoeducation program on the psychological well-being of mothers and their children, and nutritional, socioemotional, physical, and cognitive development of children by conducting a clustered-RCT on roughly 3,500 Rohingya refugee mother-child dyads. To the best of our knowledge, our study provides the first experimental evidence on improving the mental health of the Rohingya people of Myanmar through psychosocial support and its impact on early childhood development. We also provide the first experimental evidence on the impact of a psychoeducation program on the intergenerational transmission of mental health. Therefore, we contribute to the emerging literature on psycho- and non-psychotherapeutic interventions and its impact on mental well-being (Bolton et al., 2003; Rahman et al., 2008; Angelucci & Bennett, 2021; Hussam et al., 2021), behavioral change (Blattman et al., 2017), and economic decision-making and outcomes (Baranov et al., 2020).[2] Our findings also relate to the literature on the effect of parenting (Cappelen et al., 2020; Doyle, 2020) and maternal mental health (Patel et al., 2004; Rahman et al., 2008) on early childhood development. In addition, we also contribute towards the understanding of the detrimental consequences of violent conflicts (Summerfield, 2000), adverse migration experiences (Stillman et al., 2009), intimate partner violence (Fischbach & Herbert, 1997), and poverty (Ridley et al., 2020) on mental health, and how that can be improved using a rigorous psychoeducation program in low-income settings.

We have organized the rest of our paper as follows. Section 2 provides the context and section 3 provides the experimental design and details of the psychoeducation intervention. Sections 4 and 7 reports the main results and heterogeneous treatment effects respectively. We briefly conclude with the final section (8).

2 Background on the Rohingya People

> "They burnt our house and drove us out by shooting. We walked for three days through the jungle."
>
> Mohammed, a Rohingya man (UNHCR Emergencies, 2019).

> "At that moment I felt like I was already dead. I think I am only alive to tell the world about what I saw."
>
> Rajuma, a Rohingya woman (Motlagh, 2018).

The Rohingya people of Myanmar (previously Burma) are an ethnic, linguistic, and religious minority who are usually referred to as 'Bengali foreigners' by the other ethnic groups in

[2]However, light-touch and short psychotherapy programs were found to be ineffective in influencing psychological or economic outcomes in low-income settings (Haushofer et al., 2020).

their homeland. They have lived in the Rakhine State for centuries until they were stripped of their citizenship in 1982 as they did not meet citizenship requirements erected by the then Burmese government (Cheung, 2011; Mahmood et al., 2017). Rohingyas have been subject to repeated waves of persecution and forced displacement since Myanmar's independence in 1948. Around 200,000 Rohingyas fled to Bangladesh in 1978 when the Burmese military started a violent operation to screen out 'foreigners' from citizens (Cheung, 2011). Similar operations and displacement also took place after the 1991-92 elections. Likewise, in late 2012, due to increased communal violence and conflict against the Rohingya by the Rakhine residents, and later by the Burmese military, Rohingyas once again began fleeing Myanmar to their neighboring countries. A new wave of violence against the Rohingya people spurred in 2017, also known as 'ethnic cleansing' by the Burmese military, that forced the majority of Rohingyas to seek refuge in neighboring Bangladesh (Beyrer & Kamarulzaman, 2017; Bhatia et al., 2018). Since 2017, almost 1 million Rohingya people have been residing in crowded settlements in southern Bangladesh, among which 81% arrived after the 2017 incident (UNHCR Population Factsheet, 2019). This makes them one of the largest groups of stateless people in the world (Bhatia et al., 2018).

Compared to previous attempts of ethnic cleansing by the Myanmar government, the most recent attempt had the largest impact on the Rohingya population, displacing more than two-thirds of the entire 1.5 million Rohingya population living in Myanmar (Mahmood et al., 2017). According to UNHCR Population Factsheet (2019) and UNHCR Camp Profiles (2019), among the 1 million refugees currently residing in Cox's Bazar, Bangladesh, 55% are children with 41% being below the age of 11 and 18% below the age of 4. Also, 52% of the overall refugees are female. Besides, 16% of 211,383 Rohingya families residing in refugee camps are run by single mothers and 63% of families have over 4 family members. Moreover, these camps consist of 31% vulnerable families, such as separated children and families with single mothers, with at least one protection vulnerability.[3]

3 Experimental Design

3.1 The Home-Based Psychoeducation Program

The program. The international development organization BRAC Bangladesh has developed a multifaceted psychoeducation program—also known as the 'Humanitarian Play Lab'—to foster psychosocial well-being of Rohingya women who are mothers of children below 2 years as well as mental, nutritional, socioemotional, physical, and cognitive development of their children under 2 years. Specifically, experts from BRAC Institute of Education and Development and various external psychologists developed this low-cost, easy-to-implement one year program for the Rohingya mothers and children living in refugee camps in Bangladesh. It includes 44 different session modules to be delivered on a weekly basis, in a home setting. This program was developed as an urgent measure for the severely persecuted and forcibly displaced Rohingya mothers and children, with a view to implement it at-scale shortly after evaluating its impact through a randomized controlled trial.

[3]More details on demographics and camp profiles can be found in UNHCR Population Factsheet (2019) and UNHCR Camp Profiles (2019).

The psychoeducation program aims to help the Rohingya refugee mothers to cope with various mental distresses and psychological trauma that they have experienced during the 'ethnic cleansing' attempt in Myanmar and also provides a safe space for sharing (in small group sessions) their various negative and positive feelings with other Rohingya mothers. The primary aim is to ensure that mothers get a thorough understanding of the challenges they are facing, and their own coping ability and areas of strengths and weaknesses. This should allow mothers to better address their day-to-day problems, feel more in control of their feelings and actions, and to have higher internal capacity to achieve psychological well-being.

Another feature of this multifaceted program is that it allows mothers to play various culturally appropriate games with their children and other participating mothers. The program also involves advising mothers about childcare and the importance of playing with children for their healthy development. Finally, children also engage in free-play activities with age appropriate toys, such as balls, dolls, blocks, cars, stuffed animals, and so on. The program is delivered by trained Rohingya women who volunteered for the task (known as Mother Volunteers or MV) and are from the same neighborhood as the participants.[4] Thus, it was a peer-delivered intervention, provided to small groups of participating mothers at MV's home on a weekly basis, with each session lasting for 60 minutes. For the intervention, MVs were trained by mental health experts and received support (when required) by a mental health counselor and a project assistant during the intervention period.

Session procedure. Enrolled mothers in the treatment arm were required to attend a session (which is the home of a MV) along with their children on a weekly basis. Each session is divided into four steps: (i) greetings, (ii) my well-being, (iii) child's growing up, and (iv) homework. The order and description of the four steps in each session are provided below:

- **Step 1:** The first step is called ***Greetings***, which involves greeting one another and doing breathing exercises, primarily to induce relaxation among attendees. MVs also ask participants about their homework from the previous week and participants share their feedback. This step is identical in every session and runs for 15 minutes. For more details, see Figure A1.
- **Step 2:** The second step is called *"Amar Bhalo Thaka"* or ***My well-being***, which focuses on providing mental health support to mothers. The topic varied every week, which involved sharing of positive and negative feelings with other participating mothers and MV, sharing happy memories, advice on self-care (such as healthy diet, importance of sleep, nurturing hobbies, etc.), importance of communication, positive thinking and how it can help coping, acceptance and tolerance, emotional development, and so on. Mothers also played various games, such as hole tarp, bank-a-ball, etc., and participated in various art activities. During this step, children also engage in free-play activities with age appropriate toys. This step differs across sessions, where mothers receive different self-care and relaxing advice and engage in different play and art activities across sessions. This step runs for 20 minutes. For more details, see Figure A2.

[4]MVs were hired by BRAC program managers based on their level of education, fluency in Bangla and Rohingya languages, and field management skills. Priority was given to women who could also provide their home to run group sessions.

- **Step 3:** The third step is called *"Shishur Bere Otha"* or ***Child's growing up***, which focuses on parenting. Participating mothers received advice from MV on childcare, such as cleanliness, feeding, nutrition, ways to massage a baby for better sleep, etc., and the importance of playing with children and its influence on children's psychological and physical well-being. Mothers are shown how they can play with their toddlers and children with various household items, such as using pillow, handkerchief, etc. Mothers also engage in play activities with their children during this step (e.g., peekaboo, toy hunt, counting fingers, etc.). Analogous to Step 2, this step also differs across sessions, where mothers receive different childcare advice and engage in different play activities with their children across sessions. This step runs for 20 minutes. For more details, see Figure A3.
- **Step 4:** The final step is called ***Homework***, where participating mothers are assigned homework based on the weekly topic. This step runs for 5 minutes.

Randomly selected mother-child dyads in the treatment group received the psychoeducation program. In contrast, mothers in the control arm did not receive the psychoeducation program but attended an *unstructured* social gathering with other mothers (thus, there was no MV to administer psychoeducation) on a weekly basis. This allows us to disentangle the effect of psychoeducation from the effect of attending social gatherings. All sessions were conducted in the local Rohingya language. We describe our sampling and randomization in detail in section 3.2.

COVID-19 and mobile phone sessions. Due to the COVID-19 pandemic, Bangladesh went into a nationwide lockdown on March 26, 2020. Thus, after delivering 24 in-person sessions, the remaining 20 sessions were conducted over mobile phones (via basic feature phones) due to strict social distancing rules. To accommodate over-the-phone sessions, session duration and structure were carefully revised by experts from the BRAC Institute of Education and Development. Thus, various play activities, breathing exercises in groups, interaction with children during sessions, and various group discussions could not be conducted over the phone. Otherwise, session steps, topics, and the order remained unchanged. These sessions were conducted by the same MVs as in the group sessions and were implemented on a one-to-one basis. Each over-the-phone session lasted for 20 minutes. In total, 87% of enroled women from baseline had mobile phone access (same across treatment groups) and, thus, roughly 87% of treatment group mothers could participate in over-the-phone sessions. During this period, control group women did not receive any (placebo) calls or did not engage in any unstructured social gatherings. There were 20 weekly over-the-phone sessions. More details on over-the-phone sessions are provided in Appendix B.

Program Timeline. Table 1 shows the intervention timeline. Psychoeducation began in early-October 2019 and ended in late-September 2020. However, it temporarily ceased its operation in late-March 2020 due to the COVID-19 lockdown. Later, over-the-phone support replaced face-to-face sessions, which began in mid-May 2020 and ended in late-September 2020. Our baseline data collection began in late-July and ended in late-September 2019, whereas the endline data collection began in early-October and ended in late-October 2020. The endline survey was conducted over the phone due to COVID-19 restrictions. We describe the surveys

in detail in section 3.3.1.

3.2 Research design

We evaluate the home-based psychoeducation program using a cluster randomized control trial. Following arrival in Bangladesh, Rohingya people were provided with shelter in closely located refugee camps in the Cox's Bazar district. Each of these camps consists of many *Majhee* blocks (blocks hereinafter), which are clusters of many households and can be considered "neighborhoods". We use this geographic-level information, which is blocks within the camps, for randomization in our cluster-RCT study. At the time of randomization, there were over 2,000 blocks distributed across 17 refugee camps where BRAC operates (out of 34 camps in total). We randomly selected 251 blocks from the universe of over 2,000 blocks, of which 137 were assigned to the treatment (54%) and 114 were assigned to the control group (46%). Figure A4 shows a camp map and blocks therein, highlighting treatment and control blocks.

Within each block, we randomly created two groups, where each group attended a mother volunteers's (MV) home throughout a year to attend psychoeducation sessions (session hereinafter). We had a total of 226 groups in treatment and 191 in control blocks.[5] For each session, we randomly invited roughly 7 mother-child dyads on average. BRAC has a list of Rohingya households with details on their family size and demographics. From this list, project assistants and MVs randomly visited households that met the criteria, i.e., mothers with at least one child between the age of 46 days and 24 months and invited the mothers to participate in the home-based psychoeducation program and if a mother had multiple children that fall under this age category then we randomly selected one child. A total of 3,499 mother-child dyads enrolled to participate in this program. The program take-up rate was over 95% because Rohingyas residing in refugee camps are not allowed to leave their camps or be employed. Hence, mothers are mostly available to participate in such activities after finishing their daily household chores. Less than 5% could not participate because they were either responsible to take care of the elderly or they needed permission from their spouses. Only mothers in the treatment arm received our weekly psychoeducation, while mothers in the control arm participated in unstructured (or unsupervised by an MV) social gatherings that did not involve psychoeducation.

3.3 Data

3.3.1 Surveys

We collected survey data both at baseline and endline of this program (see Table 1 for the project timeline). The baseline was conducted by both female and male enumerators—trained by experts from BRAC—on a one-to-one basis (face-to-face) with participating mothers. The endline, on the other hand, was conducted over the phone due to COVID-19 restrictions. Enumerators are Bangladeshi from the Ukhiya region in the Cox's Bazar district and are fluent in the *Rohingya* language. They are highly trained with several years of survey experience. Baseline questionnaires are divided into three broad parts that cover: (i) socioeconomic background; (ii) mother's characteristics, such as demographics, migration and camp experiences,

[5]Note that blocks that were relatively small in size only had one group, instead of two.

exposure to violence both during the conflict in Myanmar and at the refugee camp, mental health conditions, current state of life, etc.; (iii) adverse life experience of children, such as demographics and mental health conditions; and age-specific cognitive skills development of children, such as communication, motor, problem-solving, and personal-social skills. Dedicated and highly trained anthropometric enumerators also measured and collected children's height and weight using appropriate scales at baseline. At endline, due to COVID-19 restrictions, trained anthropometric enumerators instructed mothers over-the-phone to measure children's height and weight in the following way: (i) height was measured by mothers using their right hand, and (ii) weight was measured by asking mothers to first weigh a one kilogram of rice packet (holding using both hands) and then lifting their children (holding using both hands) and then reporting back to the enumerator.[6]

At baseline, a total of 3,499 mothers were surveyed: 1,911 from the treatment and 1,588 from the control groups. At endline, 2,845 mothers were surveyed, 1,679 from the treatment and 1,166 from the control arm. Therefore, by the endline, roughly 19% of mothers could not be surveyed, where 13% were due to their lack of access to mobile phones and the remaining 6% was due to not answering calls despite multiple attempts by enumerators. For simplicity, we call both groups 'attritors' (no-phone contacts plus non-respondents). We discuss attrition in detail in section 3.5.

3.3.2 Outcomes

We have five mother and ten child outcomes, all pre-registered at the AEA RCT Registry. Outcomes, other than child anthropometrics, were measured by aggregating survey questions. Using these survey questions, we construct their respective indices by: (i) converting each response into a dummy, so if a question was answered on a 5-point Likert scale, then the maximum two response points were coded as 1 and the remaining three points as 0; (ii) then aggregating the dummy responses into a scale; (iii) from each response, subtracting the control group mean; and, (iv) finally dividing it by the control group standard deviation. Eventually, we get indices such that the control group has mean zero and standard deviation one. We define our outcomes below:

Mental health of mothers and children

Psychological trauma. We broadly explore psychological trauma by combining post-traumatic stress disorder (PTSD) and acute stress disorder (ASD), as their symptoms overlap.[7]

[6] Mothers used right hand and index finger lengths to measure children's height. Hand length is the length between the mid-point of the wrist's distal transverse crease and the tip of the middle finger, and index finger length is the length between the border crease of the palm and the tip of the index finger. Height of children in hand and finger lengths was then converted to centimeters following Asadujjaman et al. (2019).

[7] Trauma and stressor-related disorders emerge due to exposure to traumatic and stressful events and include post-traumatic stress disorder (PTSD), acute stress disorder (ASD), reactive attachment disorder, disinhibited social engagement disorder, and adjustment disorders (American Psychiatric Association, 2013). We only look at the impact on PTSD and ASD, the two most common disorders developed from exposure to injury, violence, or death, and leave out the latter three because reactive attachment disorder and disinhibited social engagement disorder occur mostly among children due to inadequate parenting and neglect. Moreover, we leave out adjustment disorders because the symptoms are very similar to that of depression and is also a primary driver of depressive disorder at a later stage.

According to the American Psychiatric Association (2013), the primary etiological reason for developing PTSD and ASD is experiencing traumatic events that involve exposure to actual or threatened violence (sexual, physical, verbal, etc.), serious injury, or death. An adult or a child is said to suffer from trauma if s/he experiences any combination of the following four major symptoms: (i) intrusion (such as having recurrent distressing memories or flashbacks of the traumatic event, nightmares, etc.), (ii) avoidance (such as avoiding memories or thoughts of the traumatic event or people, conversations, or activities that remind of such events), (iii) negative mood (such as unable to feel happy or satisfaction), (iv) marked alterations in arousal and reactivity (such as sleep disturbance, being easily startled, or emotional outbursts). To measure mothers' trauma through survey questions, we partially use the simplified Kessler Psychological Distress Scale (Andrews & Slade, 2001) and combine it with other survey questions that were developed following the diagnostic criteria for PTSD and ASD as laid out in American Psychiatric Association (2013). We measure children's psychological trauma through an adverse life experience survey by Dyregrov et al. (2000); Neugebauer et al. (2009).

Depression. People develop depression due to various reasons, where experiencing violence (sexual, physical, verbal, etc.) or loss of family and identity are thought to be the major drivers of depression. According to American Psychiatric Association (2013), the diagnostic criteria for a major depressive episode, or depression, is experiencing at least five of the following nine symptoms: (i) feeling sad, empty or hopeless almost every day, (ii) loss of interest in various activities, (iii) dramatic change in body weight or appetite, (iv) difficulty sleeping or insomnia, (v) feeling restless, (vi) feeling tired or fatigue, (vii) feeling worthless or guilt, (viii) unable to concentrate or make decisions, and (ix) recurrent thoughts of death, suicide, or self-harm. To measure depression, we use the Center for Epidemiologic Studies Depression Scale-20 that consists of self-reported measures of depressive symptoms (Radloff, 1977). We measure children's depressive symptoms using an adverse life experience survey by Dyregrov et al. (2000); Neugebauer et al. (2009).

Subjective well-being of mothers

We measure subjective well-being of mothers that look into their levels of happiness, hope and aspirations about the future, and their feelings of belonging to the host country in the following way:

Happiness. We explore happiness of mothers both in absolute and relative terms. We combine two survey responses that ask mothers to report their level of happiness based on their current state of lives and that relative to their friends, relatives, or neighbors.

Aspirations. We explore different aspects of future aspirations of mothers. Specifically, we measure their hopefulness about their family's future and children's formal education, possibilities of their own and husbands' employment and earnings, and the possibility of their relocation and return to their homeland.

Belongingness. People in the diasporas often struggle to 'belong' to a particular country and, hence, go through the psychological stress of searching for identity and the need to belong (Kumsa, 2006). Therefore, to measure mothers' belongingness to their new home in Bangladesh, we ask about their feelings of assurance about daily living and their feelings on

safety and security relative to their homeland.

Child development

We investigate development outcomes of children that explores different developmental progress associated with their socioemotional, cognitive, and physical development at an early age. These are communication, gross-motor, fine-motor, problem-solving, and personal-social skills. To measure these five-set of skills, we use the widely used Ages and Stages Questionnaire (ASQ-3) (Squires & Bricker, 2009), where questions are grouped into categories dedicated to assessing a specific set of skills and are also age-specific. That is, this questionnaire has different questions for 2, 4, 6, etc., months-old children.[8] All survey questions are answered by the mothers.

Communication skills: Communication skills exhibit a child's ability to perform age specific communication with mothers or others, such as making noises, chuckling and smiling, responding with sounds, constructing small sentences, following simple instructions without repetition, describe things, and so on.

Gross and fine-motor skills: Both skills define the physical performance of age-specific activities. Gross-motor skills explore the development of large body muscles and eye-hand coordination skills, such as standing, walking, jumping, throwing, catching, kicking, etc. Fine-motor skills, on the other hand, explore the development of small body muscles using wrists, fingers, toes, etc., such as holding toys, grabbing a person's finger, turning book pages, vertically arrange small boxes, etc.

Problem-solving skills: Solving simple problems such as observe and follow (hand or object) movements, drawing lines, observing daily activities, putting in and taking out stones from bottles, rearrange items, etc.

Personal-social skills: Achieving age-specific abilities about self-regulation, compliance, adaptive functioning, autonomy, interaction with people, etc. For instance, smile back if smiled at, smiling when looking at a mirror, eating by her/himself, copying what mothers do, caring for toys, etc.

Child anthropometrics

We explore children's stunting (or shortness), underweight (or lightness), and wasting (acute undernutrition) by looking at their height-for-age (HAZ), weight-for-age (WAZ), and weight-for-height (WHZ) z-scores respectively following the World Health Organization (WHO) defined guidelines (WHO, 2009).[9] According to WHO, the criterion for stunting is when $HAZ < -2$ standard deviations (i.e. 2 standard deviations/SD below the median in reference population), severe stunting is when $HAZ < -3$ SD, underweight is when $WAZ < -2$ SD, severe underweight is when $WAZ < -3$, wasting is when $WHZ < -2$, and severe wast-

[8] Questionnaires for 2 months-old children applies to 0-2 months-olds, 4 months-old children applies to 3-4 months-olds, 6 months-old children applies to 5-6 months-olds, and so on.

[9] We compute anthropometric z-scores using the 2006 WHO child growth standards. The formula to compute z-score of child i is: $z - score = [(observed\ value) - (median\ reference\ value)] \div (standard\ deviation\ of\ the\ reference\ population)$, where the median reference value is the median height or weight of all girls/boys who are i's age and the reference population is (American) children under 2 years who are chosen by WHO and National Center for Health Statistics. See WHO (2006, 2009) for more details.

ing is when $WHZ < -3$. We use both z-scores and dummy variables constructed using these cut-offs as our anthropometric outcomes.[10]

3.4 Sample characteristics and balance checks

We report the balance on observables at baseline between treatment and control groups in Table 2 and balance on baseline outcomes in Table A1. In both tables we report means and standard deviations. To derive p-values on tests of equality of means across treatment and control groups, we regress the variable of interest on the binary treatment (equals to 1 if treatment group and 0 otherwise) with camp fixed effects and standard errors clustered at the unit of randomization (which is blocks). We also report the Randomization Inference p-value (Young, 2019), which is derived from a permutation test at the level of randomization (with 1,000 replications). Balance checks in both tables are also robust to excluding camp fixed effects from the specification. Our mother and child samples are well balanced across individual and household characteristics and average differences in almost all observables are very small. Although, there are two exceptions. First, mothers in the treatment group are about 6 months older than mothers in the control group, and this difference is significant at 5% level ($p = 0.04$); and, second, children in the treatment group are 0.69 centimeters taller than their control group counterparts ($p = 0.07$). In terms of outcomes at baseline (Table A1), our samples are well balanced across all outcomes. Comparing the differences in distributions of mental health at baseline (shown in Figure A5 in Appendix A), we find that treatment group distributions are not statistically different from control group distributions using a Kolmogorov-Smirnov test (all $p > 0.10$).

Therefore, out of 35 tests, only one yields a p-value less than 0.05 and two less than 0.10. In addition, following Imbens & Wooldridge (2009) and Imbens & Rubin (2015) we also compute the normalized differences in means for all variables to show the scale-free differences.[11] The idea is that increasing the sample size can increase the t-statistic but it does not systematically affect the normalized difference. We find that out of the 35 normalized differences, 33 differences are lower than $1/8^{th}$ of the combined sample variation, one difference is below $1/3^{rd}$ (variable 'child victim of at least one camp abuse'), and another difference is below half of the combined sample variation (variable 'child's victimization (camp)'). The general rule of thumb is that if a difference exceeds one quarter, then linear regression methods are likely to be sensitive to specification changes (Imbens & Wooldridge, 2009). In any case, we control for all characteristics that differ in terms of mean or normalized differences when estimating treatment effects (see section 3.6).

3.5 Attrition

We successfully followed up on 2,845 mother-child pairs (out of 3,499) at endline, 1,679 in the treatment group (out of 1,911) and 1,166 in the control group (out of 1,588). Given the

[10] We also pre-registered mother-child relationship as an outcome (also measured at baseline) using the Ages and Stages Questionnaires: Social-Emotional (ASQSE-2) (Squires et al., 2015). However, due to limitations on the questionnaire length and interview duration at endline (during COVID-19), we could not administer ASQSE-2.

[11] We compute the normalized difference for each variable by first taking the difference in means (treatment group mean minus control group mean) and then dividing this difference by the square root of the sum of the variances.

follow-up survey was conducted in the midst of the COVID-19 pandemic, the overall attrition rate was surprisingly low at 19%. Here attrition due to not having mobile phones is 13% and non-response is 6%. However, we observe the attrition rate to be about 14 percentage points higher in the control group relative to the treatment group (T-test: $p < 0.01$). This 14pp translates to 190 women or 5% of 3,499, and as mobile phone ownership was similar across arms, this implies that most non-response was in the control group.

To check if individual and household characteristics measured at baseline predict attrition, we compare means of mothers/children that attrited to means of mothers/children that remain at endline in Table A7 in Appendix A. We find that characteristics are fairly similar, with the exception of mother's being the household head (significant at 5% level). We also observe that attrited mothers are 'marginally' newer in the camp and children are shorter in height (both significant at 10% level), but the magnitudes of these differences are very small. To further examine the differences in baseline observables, we regress attrition (equals to 1 if attrited at endline and 0 otherwise) on baseline characteristics, treatment dummy, and the interaction between baseline characteristics and the treatment dummy (see Table A8 in Appendix A). A joint F-test on the interaction coefficients yields a p-value=0.19, suggesting attrition was not differential by baseline characteristics.

Although this suggests attrition to be uncorrelated with observable characteristics, the 14 percentage points gap in response rate between treatment and control groups raises the concern that attrition may bias the treatment effects estimated in section 4. We address this concern using four different approaches, one parametric and three non-parametric, in section 4.4.

3.6 Empirical strategy

Impact on outcomes. To test the impact of the counseling program on the outcomes of mothers and children, we postulate our main empirical model as follows:

$$Y_{1ijc} = \beta_0 + \beta_1 Treat_{jc} + \beta_2 Y_{0ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \epsilon_{ijc} \tag{1}$$

where Y_{1ijc} denotes the outcome of mother/child i in block j located in camp c, measured at the endline. $Treat_{jc}$ is a binary variable that indicates the treatment status of block j in camp c; and $\boldsymbol{X}_{ijc}$ is a vector of controls, measured at the baseline, that include age, gender of child, whether mother attended school, monthly household spending, household size, number of children (or siblings for child outcomes), months spent in the refugee camp, whether the mother receives food vouchers, whether mother is the household head, whether husband is alive, whether at least one family member remains stranded in Myanmar, household's victimization based on conflict in Myanmar (an index between 0 and 1, where higher value corresponds to more exposure to violence), mother's and child's victimization based on refugee camp experience (both indices, between 0 and 1). Y_{0ijc} is the baseline analogue of the outcome. θ_c is camp fixed effects, so that the comparisons are between blocks in the same refugee camp. Error terms are assumed to be independent across blocks, but we allow it to be correlated within the same block (i.e., the standard errors are clustered at the block level, which is our unit of randomization). As outcomes other than anthropometrics are indices with the control group having mean 0 and standard deviation 1, our estimated β_1 for such outcomes determines where the

mean of the treatment group lies in the distribution of the control group in terms of standard deviations (Kling et al., 2007). Since trauma and depression indices are based on 'negative' feelings, negative $\hat{\beta}_1$ correspond to improvement in mental health. For the remaining outcomes, positive coefficients correspond to more favorable outcomes. We estimate equation 1 using OLS.

Inference. First, we cluster standard errors at the unit of our randomization, which is by blocks. Second, even though the number of clusters per arm is somewhat large (more than 110 clusters in each arm), for robustness, we also compute p-values using randomization-based inference (RI) with randomization permuted at the cluster level (Young, 2019). For this, we use 1,000 replications.[12] In regression tables that report treatment effect estimates, we also report the Young (2019) RI p-values. Results reported in the following section are largely robust to using this method.

Correction for multiple hypotheses testing. With many outcomes of interest (15 outcomes), it is crucial to correct p-values for each outcome that we test. To address this issue, we check the robustness of our results using the List-Shaikh-Xu procedure that uses bootstrapping (with 3,000 replications) to account for joint correlation across different tests and then controls the probability of making any type-I error (or the familywise error rate (FWER)) (List et al., 2019). In each regression table, where we report the treatment effects, we also report the FWER-adjusted p-values for each test. We also check the robustness of our results using the Westfall-Young adjustment (Westfall & Young, 1993). Our conclusions are largely consistent using both methods.

4 Results

We divide this section into four parts. First, we present results on the impact of the intervention on all outcomes (section 4.1). Second, we present results on whether the intervention affected transmission of mental health from mothers to children (section 5). Third, we investigate the channels through which children's mental health were improved (section 6.2). Finally, we explore the impact of differential attrition on the treatment effect estimates (section 4.4).

4.1 Impact of the intervention on mothers

Mental health. Figure 1 and Table 3 report the impact of the intervention on mother's mental health (Panel A1) and subjective well-beings, such as happiness, aspirations, and belongingness (Panel A2). Column 1 reports treatment effects without controlling for baseline controls and column 2 reports estimates with the full set of controls. Since results with and without controlling for baseline characteristics are similar, we focus our discussions below only based on estimates reported in column 2.

We find that the intervention has significantly improved mental health of *Rohingya* mothers. Specifically, mothers that received mental health counseling experienced a 0.23 standard deviation (SD) reduction in trauma ($p < 0.01$) and 0.14 SD reduction in depression ($p < 0.01$)

[12] Young (2019) suggests that draws beyond 2,000 make little to no difference to p-values. Our conclusions do not change if we use 2,000 replications.

relative to mothers in the control group that did not receive any mental health counseling (Panel A1, Table 3). When we explore the impact of the intervention only among mothers that were identified as traumatized (1,557 out of 3,499 mothers) and depressed (645 out of 3,499 mothers) at baseline, we find that the reduction in trauma is 0.26 SD among the traumatized, which is slightly higher than the aggregate effect of 0.23 SD. However, the reduction in depression among the depressed is sizable at 0.29 SD (a twofold improvement) relative to depressed mothers in the control group. In comparison to short-run impacts of other mental health interventions in developing countries (Baranov et al., 2020; Vlassopoulos et al., 2021), our estimated impacts are relatively smaller. One possible reason is that participants in the control group in our study also participated in social gatherings on a weekly basis (which is mostly muted in existing studies), and as social interaction has a positive impact on people's mental well-being (Nezlek et al., 1994), it is possible that improvements in mental well-being of our control group participants must have contracted the treatment effect. On the other hand, our effect sizes are largely comparable to the effect sizes of a recent non-therapeutic study conducted on the Rohingya refugees in Bangladesh (Hussam et al., 2021).

Subjective well-being. We then consider outcomes related to mother's subjective well-being in terms of happiness, aspirations for the future, and belongingness (Panel A2). We find that happiness and belongingness of mothers in the treatment group increased by 0.12 SD and 0.18 SD respectively relative to mothers in the control group. However, in terms of aspirations, treatment effect is muted. We also illustrate these treatment effect estimates in Figure 1, where we show where the mean of the treatment group lies in the distribution of the control group in terms of SD units. Under each 'pooled' results, we also present results by child's gender (graph A) and by exposure to violence during the conflict in Myanmar (graph B). We discuss these heterogeneous treatment effects in detail in section 7.

Are the mentally unhealthy catching up to the healthy? Next, in Panel A1, Table 4, we examine whether the mentally unhealthy mothers in the treatment group are catching up to the mentally healthy mothers in the control group in terms of mental and subjective well-beings. We find that our intervention was successful at aligning the depression severity of the treated mothers that were depressed at baseline with the depression severity of the control mothers that were healthy at baseline (suggested by the statistically insignificant coefficients), implying that the mentally unhealthy in the treatment group successfully caught up to the mentally healthy in terms of depression. On the other hand, in terms of trauma, we find that the mentally unhealthy mothers in the treatment group, in fact, surpassed the mentally healthy mothers in the control group by 0.20 SD, which is sizable and statistically significant at 1% level. Finally, in terms of subjective well-being (Panel A2), the mentally unhealthy mothers in the treatment group also surpassed the subjective well-being of the mentally healthy mothers in the control group. Thus, the psychoeducation treatment was very effective at not only lifting mothers from mental distress but also making their mental health better than those who were considered mentally healthy at baseline.

4.2 Impact of the intervention on children

Mental health. Figure 1 and Table 3 also report treatment effects on child-outcomes (Panels B1 and B2). We find that children experienced an improvement in both psychoeducation trauma and depression severity (Panel B1)—a 0.10 SD reduction in trauma ($p < 0.10$) and a 0.12 SD reduction in depression ($p < 0.05$) relative to children in the control group. When we examine treatment effects only among the traumatized and depressed (at baseline), we find conclusions analogous to that among mothers. That is, the effect on trauma among the traumatized is 0.13 SD ($p < 0.10$), which is slightly larger than the aggregate impact of 0.1 SD. On the other hand, the effect on depression among the depressed is twofold larger (a 0.24 SD reduction) than the aggregate effect ($p < 0.05$). In section 9, we conduct a causal mediation analysis to explore how much of the children's mental health effects are directly associated with the intervention and how much of it is mediated via improvements in mothers' mental health.

Socioemotional, physical, and cognitive development. Next, we examine the impact on children's developmental trajectories (Panel B2). We find that the intervention significantly improved communications skills of children by 0.23 SD ($p < 0.01$), gross-motor skills by 0.18 SD ($p < 0.01$), problem-solving skills by 0.18 SD ($p < 0.01$), and social skills by 0.13 SD ($p < 0.10$) relative to children in the control group. While the impact on the former three domains are statistically significant at 1% level, impact on child's social skills is rather weak and only marginally significant at 10% level. In terms of fine-motor skills, we do not find any statistically significant treatment effect.

Anthropometrics. In terms of child malnutrition (Table 5), we estimate treatment effects on children's underweight, stunting, and wasting. First, we find that treatment improved treated children's weight-for-age z-score (WAZ) by 0.64 SD, which is equivalent to increasing the weight of treated children by 635 grams relative to children in the control group (column 3, Panel A). Also, we observe a significant reduction in underweight by 7pp (which is 10% relative to the control group) and a reduction in severe underweight by 16pp (which is 26% relative to the control group), implying children at the lower and extreme lower tails of the distribution benefited greatly by the intervention.

Second, we consider height-for-age z-scores (HAZ) that also measures malnutrition but focuses on skeletal growth retardation of children. We find that our intervention was successful at increasing the HAZ of treated children by 0.52 SD, which is equivalent to 1.58 centimeters (column 3, Panel B). At the lower and extreme lower tails of the distribution, we observe stunting and extreme stunting fell by 7pp (or 10%) and 13pp (or 22%), respectively. Finally, we examine impacts on children's prevalence of wasting, measured by weight-for-height z-scores (WHZ). Both improvements in weight and height are reflected on the incidence of wasting, where we observe WHZ of treated children to increase by 0.51 SD. Moreover, treated children's prevalence of wasting reduced by 8pp and severe wasting reduced by 13pp relative to children in the control group.

These differences that we observe do not differ by children's gender (columns 4-6), suggesting both female and male children experienced reductions in nutritional deprivation relatively equally.

4.3 Robustness checks

We examine the robustness of our results in three ways. First, we examine whether mothers' tendency to give socially desirable response bias our treatment effects. Second, we check whether mental health of mothers might be affecting their judgments on child development, which as a result might be biasing survey response on child development. Finally, we check whether 'highly influential' observations might be driving our results.

Experimenter demand effects. One key concern with self-reported outcomes is that respondents often have the tendency to provide responses to survey questions that might be deemed favorable by surveyors (social desirability bias), and receiving some 'treatment' from surveyors or their employers might trigger such behavior (experimenter demand effects). For instance, in our context, treated respondents that received psychoeducation for a year might feel more inclined to provide favorable response to enumerators relative to control group respondents. However, in this study, control group participants also participated in social gatherings (pre-pandemic) organized and invited by BRAC Bangladesh. This provides some reassurance that experimenter demand effects might also be present among the control group participants (thus, possibly constant across treatment arms). However, our program objectives were very salient to women in the treatment group and, hence, experimenter demand effects might remain a concern.

To carefully address this issue, we measured our respondent's general tendency to provide socially desirable responses using a 13-item Marlowe-Crowne scale (Crowne & Marlowe, 1960; Dhar et al., 2021). This scale was developed by psychologists and has been validated in various contexts. The questionnaire asks whether respondents have various too-good-to-be-true personality traits—such as whether respondents are excellent listeners or never hurting anyone's feelings on purpose—to create a social desirability bias or SDB scale ($0 \leq SDB \leq 13$). The higher the scale, the higher is a respondent's tendency to give socially desirable answers. Using this scale, we carry out a heterogeneity analysis to check whether people that score higher on the Marlowe-Crowne scale are more likely to experience stronger treatment effects (evidence for experimenter demand effects). We report this result in Table 7, where we find that the coefficients on the interaction term never reach statistical significant at conventional levels. More importantly, among the respondents with low SDB score (hence less likely to give socially desirable responses), we find that our treatment effects remain sizable and statistically significant at conventional levels (coefficients on 'Treatment'). This robustness check, therefore, suggests that our main results are less likely to be a product of experimenter demand effects. Moreover, strong results on child anthropocentric outcomes, which are relatively more objective measures as they are factual and can be validated by BRAC in the future, corroborates this claim.

Judgment of mothers. Our next concern is whether results on child development—which are carefully observed and then reported by mothers—are influenced by mothers' judgment. It is widely known that being mentally unwell can impair a person's attention to detail or their short-term memory, which can eventually affect their judgment. For instance, a depressed mother might not have the mental strength to carefully observe their child or remember various incidents that are strong indicators for child development, while a non-depressed might not face

such problems. While this is potentially an important issue, we are confident that our results are unlikely to be a product of such 'judgment bias'. First, enumerators from BRAC are highly trained with several years of experience surveying respondents from low-income households. They were also specifically trained to be very patient with our respondents to allow them ample time to recall and carefully answer to questions. Second, many of the child development questions were validated by enumerators during the interview, for example, how quickly child grabs mother's finger, whether child follows a toy when moved around, can jump, responds to mother's calling, arrange toys vertically/horizontally or beads in a string, etc.[13] Therefore, we can rule out the concerns of memory or attention affecting mothers' answers to child questionnaires. Finally, to empirically address this concern, we re-estimate the treatment effects reported in Table 3 by excluding mothers who experienced an *improvement/change* in their mental health. The main assumption is that mothers that remained mentally the same at endline as they were at baseline (i.e., depressed at baseline remained depressed at endline or non-depressed at baseline remained non-depressed at endline), their attention to details and, hence, judgment should remain fairly constant. If among this sample we observe statistically significant treatment effects on child development outcomes, then mothers' bias in judgment must not be explaining our findings. We report these conservative estimates in Table A3, which shows that our main results on child development outcomes are robust to such extreme adjustments.

Highly influential observations. To check if our results are driven by some 'highly influential' observations (or outliers) that might be pulling the regression line towards them, we compute the 'dfbeta' for a particular observation. The idea is to compute the difference between the regression coefficient for a variable with that observation and the coefficient by dropping that observation, and then scale it by the standard error of the coefficient computed without that observation (Belsley et al., 2005). We then use the suggested cut-off of $2/\sqrt{n}$, where n is the number of observations, to drop observations for which the 'dfbeta' value is above this cut-off, since such observations tend to be overly influential on the estimated coefficient. We find that less than 1% of observations at endline requires dropping in all specifications, implying the presence of 'highly influential' observations is very small in our data. Following dropping such observations, our results remain robust and sizeable throughout, suggesting estimates are not overly sensitive to the exclusion of influential observations.

4.4 Differential attrition and treatment effects

As highlighted in section 3.5, there is significantly higher attrition in the control group relative to the treatment group (T-test: $p < 0.01$). Thus, to check whether differential attrition might have biased our estimated treatment effects in sections 4.1 and 4.2, we use four different approaches, one parametric and three non-parametric. First, we use inverse probability weighting (IPW) to estimate the treatment effects. For this, we use weights from the predicted probability of being in the endline sample based on baseline characteristics. These attrition-adjusted estimates are almost identical to the unadjusted estimates. We report the unadjusted (column 1) and IPW-adjusted (column 2) treatment effects in Table 10. Second, following Lee

[13] Since endline was conducted over-the-phone, mothers were asked to check these during interviews while the enumerator was on the line.

(2009), we conduct a trimming bounds analysis. For this, outcomes are first sorted from better to worse within treatment and control groups, then trims the sample from above and below in the treatment group (since 'excess observations' are in the treatment arm) to get lower and upper bounds. Excluding the excess observations from the treatment group results in equal sample sizes in both treatment and control groups. Our conclusions remain largely consistent with Lee (2009) bounds (column 3, Table 10). Furthermore, we use two covariates (duration in the camp and monthly household spending) that have some explanatory power for attrition (see Table A8 in Appendix A) to tighten the treatment-effect bounds. Results remain robust to using tightened bounds (column 4, Table 10).

Third, following Kling et al. (2007), we add 0.05, 0.10, and 0.025 standard deviations to the means of mental health outcomes to create a lower bound (while we subtract the same amounts for all other outcomes). These additional bounds analysis is presented in Table A9, where columns 6-9 reports estimates by adding/subtracting certain standard deviations to means. These three bounds show that our main results would hold even if the mental health (well-being and child development) outcomes of the attrited sample in the treatment group were 0.25 standard deviations higher (lower) than the control group means. Recall that, for mental health (well-being and child development), positive (negative) SD in the treatment group corresponds to unfavorable outcomes.

Finally, although based on extreme assumptions about attrition, we follow Horowitz & Manski (2000); Karlan & Valdivia (2011) to create lower and upper bounds. Horowitz & Manski (2000) yields very wide bounds due to imputing extreme values and is often suitable when outcomes are discrete and attrition is very low (Ozler, 2017). We conduct bounds analysis following Horowitz & Manski (2000) nevertheless, where we impute missing information on the basis of minimal and maximal possible values. For instance, the lower (upper) bound was obtained by imputing missing data with the minimum (maximum) value in the observed treatment distribution to attritors in the treatment group and maximum (minimum) value in the observed control distribution to attritors in the control group. In a similar manner, instead of imputing minimal and maximal values, we impute missing data with the mean value of the lowest (highest) 10% observations in the observed treatment distribution to attritors in the treatment group and highest (lowest) 10% observations in the observed control distribution to attritors in the control group for the lower (upper) bound. We report treatment effects using these bounds in columns 2-5 in Table A9. Analogous to results in Karlan & Valdivia (2011), we see that these bounds are indeed very wide due to imputing extreme values.

In summary, although we observe some degree of sensitivity while incorporating extreme bounds, our estimated treatment effects are not sensitive to trimming observations from above and below or to imputing missing information with upto 0.25 standard deviations in the treatment group. According to Table 3, column 2, the largest effect size in standard deviations unit is for trauma, which is -0.23 or 0.23 SD below the control group mean (recall negative coefficient implies improvement in mental health). Thus, imputing attrited sample in the treatment group with +0.25 SD and that in the control group with 0—implying attrited mothers in the treatment group become worse-off than attrited mothers in the control group following intervention—only changes the effect size by roughly 0.03 SD (from -0.23 to -0.20). This suggests that even with cases where the attrited sample in the treatment group experience a large negative impact of

the intervention, our main results should remain similar.

5 Intergenerational transmission of mental health

Correlation. As children under 2 years spend most of their time at home with their mothers, we hypothesized that mother-child mental health conditions would be strongly correlated. Thus, we test the transmission of mental health from mothers to children using measures of trauma and depression by looking at correlations in the spirit of Dohmen et al. (2012). Thus, while discussing correlations we do not claim the relationship to be causal. To investigate this, both at baseline and endline, we estimate the following regression using OLS:

$$y_{ijc} = \phi_0 + \phi_1 Y_{ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \sigma_{ijc} \tag{2}$$

where y_{ijc} denotes the outcome (trauma or depression) of a child of mother i in block j located in camp c. Similarly, Y_{ijc} is either trauma or depression of mothers i in block j located in camp c. Here, we test the null that $\phi_1 = 0$.

In Table A5, we report the correlation coefficients using data collected at both baseline (Panel A) and endline (Panel B). In terms of intergenerational correlation in trauma and depression, we find that the coefficients are positive and statistically significant (columns 1 and 4) (both $p < 0.01$), implying mental health of children tends to be similar to the mental health of their mothers. Figure A5 in Appendix A also illustrates similarities in mental health between mothers and children when graphs are vertically compared (A1 versus B1 and A2 versus B2). In fact, correlation coefficients became larger in Panel B (Table A5), suggesting our intervention must have made the mental health of children to be more aligned with that of their mothers. When we look at correlations by child's gender, we see that mental health of mothers are positively correlated with both female and male children (columns 2-3 and 5-6, all $p < 0.01$). Also, coefficients on the interaction between child's gender and mother's mental health are not statistically different than zero, suggesting mental health correlations do not differ by child's gender.

Causal impact. We then investigate how strongly mothers transmitted trauma and depression to their children, or vice versa, following the intervention. We claim this relationship to be causal as we exploit the variation caused by randomly assigning blocks to either treatment or control arms. To check the impact of the program on the transmission of mental health from mothers to children, we estimate the following equation using OLS:

$$\triangle_{1ijc} = \kappa_0 + \kappa_1 Treat_{jc} + \triangle_{0ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \psi_{ijc} \tag{3}$$

where $\triangle_{1ijc} = |Y_{ijc} - y_{ijc}|$ is the absolute difference in mental health (trauma or depression) between mothers (Y_{ijc}) and children (y_{ijc}) at the endline, and $\triangle_{0ijc}$ is the baseline analogue of the outcome. In this specification, if κ_1 is negative and significant, then it means the program narrowed the mental health gap between mothers and children and, thus, will imply a strong transmission of mental health from mothers to children following the intervention. In other words, it be evidence that children's mental health became more aligned with the mental health of their mothers due to the intervention.

Table 6 reports treatment effects on the mother-child mental health gap. We find that the intervention reduced the difference in trauma between mothers and children (negative and statistically significant coefficient on the treatment indicator) but we do not observe any statistically significant impact of the intervention on the difference in depression (column 2). In Table 3, we saw that the intervention improved trauma of both mothers and children, but now Table 6 reports how the intervention helped align the mental health of children to that of their mothers. We observe treatment effects on both psychological trauma and depression gaps between mothers and children, where all coefficients are statistically significant and negative (Panels A1 and B1).

When we disaggregate the impact to check if the treatment effect is similar across mothers with female and male children, we do not find any evidence of such heterogeneity. That is, the intervention reduced the difference in both trauma and depression between mothers and their sons and daughters (Panels A2 and B2). To check if coefficients reported under Panel A2 (in column 2) statistically differ, we interact child's gender with the treatment dummy and find that the coefficients on this interaction term fails to reach statistical significance at conventional levels ($p > 0.10$).

This suggests that interventions targeting the mental well-being of mothers can be an important stepping stone to developing psychological resilience among their children. This is very important in contexts where psychosocial support facilities for small children are scarce and unavailable.

6 Mechanisms

While results in section 5 provide some indications that children's mental health might have improved via mothers' mental well-being, there might be various other channels through which the outcomes of this study were affected. Thus, the purpose of this section is to explore some potential mechanisms, which we do in two steps. First, we examine the direct effect of our mutifaceted psychoeducation intervention on several intermediate outcomes (or potential mediators), and then we use a causal mediation analysis to examine how much of the treatment effects on child outcomes are mediated through mothers' mental well-being and how much of the effects are due the intervention itself (or any unobservable mediators).

6.1 Effect on intermediate outcomes

Mothers. Mothers mental health outcomes might have been affected through: (i) mothers' improvement in physical health, as several advise given in the "my well-being" part focus on healthy diet, exercise, and various physical activities that could have improved mothers' physical health; (ii) relationship with husbands, as various advice on communication or sharing might have improved relationships or spending an hour away from home on a weekly basis might have deteriorated relationships with husbands; (iii) seeking more help for household chores, as less burden from household chores might be affecting mental well-being; and, (iv) staying connected with friends and family during COVID-19 lockdown, which might have improved the mental health of mothers. We explore these four channels in Panel A, Table 8. However, we do not

find any statistically significant evidence of these four factors being plausible mechanisms. In other words, the program has possibly improved the mental health of mothers directly.

Children. We next consider several potential mechanisms for improvements in child development outcomes: (i) mothers' and fathers' time input on children, as the intervention might have encouraged both mothers and fathers to spend more quality time with their children; (ii) until what age mothers breastfed their children, as more caring mothers might have breastfed their children longer; (iii) number of times children are fed per day; (iv) reduced negative parenting (such as scolding, beating, etc.); (v) asking others to spend time with children, as spending more supervised time with other adults from the household implies more daily care for children; (vi) less indoor smoking by fathers, as mothers might have become more careful about children's health and, thus, asking fathers to smoke outdoors; and, (vii) being more careful about children walking/playing barefoot, which prevents hookworm infections and exposure to various bacterial and fungal organisms. Panel B in Table 8 reports these results. We find strong support for mothers' time input being a potential mediator, where treated mothers started spending about 1.5 hours everyday on their children ($p < 0.01$); also, this effect does not differ by child's gender, suggesting mothers spend similar time on girl and boy children. We also find that treated mothers are less likely to allow their children to play or walk barefoot ($p < 0.05$) and engage in negative parenting ($p < 0.10$), suggesting improvements in mothers' health behaviors toward their children are other potential mechanisms behind their children's development. While we do not see any gender bias in negative parenting, we observe that mothers are only careful about their sons' walking or playing barefoot but not their daughters, suggesting some gender bias in mothers' parenting. However, fathers' time input on children, breastfeeding time, feeding frequency, asking others to spend time on children, and discourage fathers to smoke indoors, are not possible channels of impact.

Given mothers' increased time-intensive investment on their children per day, we next explore in section 6.2 how much of the mothers' mental health improvements is responsible for children's mental health and development improvements by using a formal causal mediation framework.

6.2 Causal mediation analysis (TO BE UPDATED)

The two main steps of each counseling session (explained in section 3.1) are: (i) mothers receive mental health advice and support to improve their own mental health conditions; and, (ii) mothers receive advice and directions on childcare and play-based parenting. Hence, we expect the mechanisms through which children would benefit from the program are through the program directly (*direct effect*) and via improving the mental health of mothers (as improved mental health of mothers might be transmitted to their children - *mediation effect*). Therefore, we conduct a formal mediation analysis using the approach by Imai et al. (2010) to understand the channels through which children's mental health were affected by the intervention. To do so, we use the mental health outcomes of mothers as our mediators. Since both outcome and mediator variables are continuous (indices), to estimate the average causal mediation effect (ACME), we first estimate model 1 and then estimate,

$$y_{1ijc} = \alpha + \tau Treat_{jc} + \pi Y_{1ijc} + \lambda_2 y_{0ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \omega_{ijc} \tag{4}$$

where y_{1ijc} is the mental health outcome of children measured at the endline (and y_{0ijc} is children's outcome measured at the baseline); Y_{1ijc} is the mental health outcome of mothers measured at the endline (the mediator). Following Imai et al. (2010), after estimating equation 1 and then equation 4, we take the products of β_1 from model 1 and π from model 4 to get the mediation effect, $\beta_1\pi$. Since we have two mediators and two mental health outcomes of children, we estimate model 4 four times, twice for children's trauma (with trauma and again with depression of mothers as mediators) and twice for children's depression (with trauma and again with depression of mothers as mediators), to get the four mediation effects. In model 4, τ is the direct effect of the counseling program on children's outcomes. In all specifications, error terms are clustered at the block level.

Table 9 reports the average direct and mediation (ACME) effects (columns 1 and 3) and the corresponding 95% confidence intervals (columns 2 and 4). With trauma (Panel A) and depression (Panel B) of mothers as mediators, we find that the average effects of the treatment on children's trauma that operates through mothers' trauma is -0.06 and mother's depression is -0.02, both statistically significant at 1% and 5% level respectively. Similarly, we also find that the average effects of the treatment on children's depression that operates through mothers' trauma is -0.04 ($p < 0.01$) and mother's depression is -0.02 ($p < 0.05$). In contrast, estimates of the direct effects are not statistically different from zero, suggesting the intervention improved the mental health of children through improving their mothers' mental health. To examine if subjective well-being of mothers played any role in improving mental health of children, we repeat the analysis but with happiness (Panel C), aspirations (Panel D), and belongingness (Panel E) as mediators. We find that only mothers' belongingness played a strong role, while happiness and aspirations of mothers were not effective channels.

This analysis is carried out under the assumption of sequential ignorability. That is, we assumed that the error term from model 1 is uncorrelated with the error term from the mediation model (4). Thus, a violation of this assumption can lead to a correlation between the two errors and subsequently bias the estimated ACMEs reported in Table 9. To adjust this, we carry out a sensitivity analysis that computes the correlation between the error terms (ρ) and then computes the true values of ACME for given values of ρ. We see that ρ needs to be approximately between 0.15-0.24 for the ACMEs reported in Panels A and B to be zero. With $\rho = 0$, all four ACMEs in Panels A and B remain fairly stable and statistically significant at 5% level, suggesting estimated ACMEs in Table 9 are robust.

7 Heterogeneous treatment effects

To estimate whether treatment effects vary by children's gender and by household's exposure to violence during the conflict, we estimate the following interaction model:

$$Y_{1ijc} = \beta_0 + \beta_1 Treat_{jc} + \beta_2 G_{ijc} + \beta_3 Treat_{jc} \times G_{ijc} + \beta_4 Y_{0ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \epsilon_{ijc} \quad (5)$$

where G_{ijc} is either children's gender (an indicator for male) or an indicator for high exposure to violence during the conflict in Myanmar. To construct the violence exposure dummy, we use the household victimization index (see Panel A, Table **??**) and then use the median cut-off.

That is, the violence exposure dummy is equal to 1 if the household victimization index is above the median value and 0 otherwise.

We did not formulate any specific hypothesis with regards to how this intervention might affect mental health of mothers with male versus female children under 2 years. It is possible that the treatment had stronger effects on the mental health of mothers with sons as parents tend to be happier and optimistic in general if they have sons versus daughters (Raley & Bianchi, 2006). Also, prevalence of son-preference might induce mothers to be more attentive and engaging during counseling sessions, which might have affected outcomes differently. Similarly, mothers/children from households that were exposed to more violent conflict in Myanmar might have been strongly affected by our intervention relative to mothers/children from households that were exposed to less violent conflicts, as recurring memories from traumatic events should be more common among the highly exposed. Since we cannot claim violence exposure of households to be entirely exogenous, we do not claim any statistically significant heterogeneous effects by violence exposure as causal.

Table 11 presents the estimated effects. Column 1 reports the pooled effects (same as column 2 in Table 3), while columns 2 and 3 disaggregate the effects by child's gender. Column 4 reports the coefficient on the interaction between child's gender and the treatment indicator, showing the difference-in-differences. We find that, on most occasions, treatment effects appear to be relatively larger among women with male children (Panels A1 and A2). Whereas, in terms of children's mental health (Panel B1) and development (Panel B2) outcomes, female children appear to have benefited more than male children. However, differences between effects reported in column 2 and column 3 are not statistically significant at conventional levels, as suggested by all insignificant coefficients in column 4. Thus, we do not find any evidence for heterogeneous treatment effects by child's gender.

We now turn to examining heterogeneity by violence exposure. Column 5 reports estimates among the highly exposed individuals while column 6 reports estimates among the least exposed. Column 7 reports the coefficients on the interaction term. In terms of mental health outcomes (Panels A1 and B1), we find that treatment effects are larger among the highly exposed relative to the least exposed, and these differences are only marginally significant for mothers' mental health. This suggests that the improvement in mental health of highly exposed mothers are larger then the improvement observed among mothers that had low exposure to violent conflict. Surprisingly, we also find that treatment effect on mother's aspirations are negative and vary by violence exposure, where aspirations of mothers that had high exposure to violent conflict deteriorated relatively more following the intervention. However, this difference is marginally significant ($p < 0.10$). In terms of children's development outcomes, we do not find any evidence for heterogeneity by violence exposure.

We also conduct additional heterogeneity analysis by experiences of abuse in the refugee camp and by the duration in the camp. These results are reported in Table A6 in Appendix A. We find that children that did not encounter any camp abuse experienced a significant improvement in problem-solving skills than children that encountered at least one camp abuse (column 4, Panel B2). In addition, mothers that are relatively newcomers in the camp experienced larger improvements in psychological trauma than mothers who have been living in the camp for more than two years (column 7, Panel A1). In terms of remaining outcomes, we do not observe any

heterogeneity.

8 Concluding remarks

Exposure to violence, forced displacement induced by violent conflict, and adverse migration experiences can have significant impact on the mental health of vulnerable refugees, such as women and child refugees, which might consequently disrupt parenting by mothers of young children and impair physical, emotional, and intellectual development of child refugees. Therefore, to help mothers overcome mental health problems and improve parenting practices, BRAC provided weekly psychosocial support to displaced *Rohingya* mothers with children under 2 years. After a year of rigorous counseling through trained para-counselors, mental health of mothers and children improved significantly relative to mental health of mothers and children that did not receive any counseling. Furthermore, this intervention improved children's communication, problem-solving, social, and gross-motor skills. In addition, mothers reported to be happier and expressed higher sense of belonging to the host community following the intervention. These findings demonstrate the importance of providing psychosocial support to displaced mothers of young children to treat and heal psychological trauma and depression, which can have strong positive impacts on their children's cognitive development and growth.

References

Adhvaryu, A., Fenske, J., & Nyshadham, A. (2019). Early life circumstance and adult mental health. *Journal of Political Economy*, *127*(4), 000–000.

Ahmed, B., Orcutt, M., Sammonds, P., Burns, R., Issa, R., Abubakar, I., & Devakumar, D. (2018). Humanitarian disaster for rohingya refugees: impending natural hazards and worsening public health crises. *The Lancet Global Health*, *6*(5), e487–e488.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Washington DC.

American Psychological Association (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, *48*(1).

Andrews, G. & Slade, T. (2001). Interpreting scores on the kessler psychological distress scale (k10). *Australian and New Zealand Journal of Public Health*, *25*(6), 494–497.

Angelucci, M. & Bennett, D. (2021). The economic impact of depression treatment in india. *IZA Discussion Paper*.

Asadujjaman, M., Molla, M. B. A., & Al Noman, S. N. (2019). Stature estimation from hand anthropometric measurements in bangladeshi population. *Journal of Forensic and Legal Medicine*, *65*, 86–91.

Baranov, V., Bhalotra, S., Biroli, P., & Maselko, J. (2020). Maternal depression, women's empowerment, and parental investment: Evidence from a randomized controlled trial. *American Economic Review*, *110*(3), 824–59.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*, volume 571. John Wiley & Sons.

Beyrer, C. & Kamarulzaman, A. (2017). Ethnic cleansing in myanmar: the rohingya crisis and human rights. *The Lancet*, *390*(10102), 1570–1573.

Bhatia, A., Mahmud, A., Fuller, A., Shin, R., Rahman, A., Shatil, T., Sultana, M., Morshed, K. M., Leaning, J., & Balsari, S. (2018). The rohingya in cox's bazar: when the stateless seek refuge. *Health and Human Rights*, *20*(2), 105.

Blattman, C., Jamison, J. C., & Sheridan, M. (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic Review*, *107*(4), 1165–1206.

Bolton, P., Bass, J., Neugebauer, R., Verdeli, H., Clougherty, K. F., Wickramaratne, P., Speelman, L., Ndogoni, L., & Weissman, M. (2003). Group interpersonal psychotherapy for depression in rural uganda: a randomized controlled trial. *JAMA*, *289*(23), 3117–3124.

Cappelen, A., List, J., Samek, A., & Tungodden, B. (2020). The effect of early-childhood education on social preferences. *Journal of Political Economy*, *128*(7), 2739–2758.

Cheung, S. (2011). Migration Control and the Solutions Impasse in South and Southeast Asia: Implications from the Rohingya Experience. *Journal of Refugee Studies*, *25*(1), 50–70.

Collins, P. Y., Patel, V., Joestl, S. S., March, D., Insel, T. R., Daar, A. S., Bordin, I. A., Costello, E. J., Durkin, M., Fairburn, C., et al. (2011). Grand challenges in global mental health. *Nature*, *475*(7354), 27.

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349.

Currie, J. (2009). Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature*, *47*(1), 87–122.

Currie, J. & Stabile, M. (2006). Child mental health and human capital accumulation: the case of adhd. *Journal of Health Economics*, *25*(6), 1094–1118.

Dhar, D., Jain, T., & Jayachandran, S. (2021). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review (forthcoming)*.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *Review of Economic Studies*, *79*(2), 645–677.

Doyle, O. (2020). The first 2,000 days and child skills. *Journal of Political Economy*, *128*(6), 2067–2122.

Dyregrov, A., Gupta, L., Gjestad, R., & Mukanoheli, E. (2000). Trauma exposure and psychological reactions to genocide among rwandan children. *Journal of Traumatic Stress*, *13*(1), 3–21.

Fazel, M., Wheeler, J., & Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. *The Lancet*, *365*(9467), 1309–1314.

Fischbach, R. L. & Herbert, B. (1997). Domestic violence and mental health: correlates and conundrums within and across cultures. *Social Science & Medicine*, *45*(8), 1161–1176.

Haushofer, J., Mudida, R., & Shapiro, J. (2020). The comparative impact of cash transfers and psychotherapy on psychological and economic well-being.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052–86.

Heckman, J. J. & Pinto, R. (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric Reviews*, *34*(1-2), 6–31.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Herman, J. L. (2015). *Trauma and recovery: The aftermath of violence–from domestic abuse to political terror*. Basic Books, New York.

Horowitz, J. L. & Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, *95*(449), 77–84.

Hossain, A., Ahmed, S., Shahjalal, M., & Ahsan, G. U. (2019). Health risks of rohingya children in bangladesh: 2 years on. *The Lancet*, *394*(10207), 1413–1414.

Hussam, R. N., Kelley, E. M., Lane, G. V., & Zahra, F. T. (2021). The psychosocial value of employment. *NBER Working Paper*.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 51–71.

Imbens, G. W. & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Imbens, G. W. & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

Islam, M. M. & Yunus, M. Y. (2020). Rohingya refugees at high risk of covid-19 in bangladesh. *The Lancet Global Health*, *8*(8), e993–e994.

Karlan, D. & Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, *93*(2), 510–527.

Kessler, R. C. & Frank, R. G. (1997). The impact of psychiatric disorders on work loss days. *Psychological Medicine*, *27*(4), 861–873.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119.

Kumsa, M. K. (2006). 'No! I'm not a refugee!'The poetics of be-longing among young Oromos in Toronto. *Journal of Refugee Studies*, *19*(2), 230–255.

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

Lindert, J., von Ehrenstein, O. S., Priebe, S., Mielck, A., & Brähler, E. (2009). Depression and anxiety in labor migrants and refugees–a systematic review and meta-analysis. *Social Science & Medicine*, *69*(2), 246–257.

List, J. A., Shaikh, A. M., & Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, *22*(4), 773–793.

Lukens, E. P. & McFarlane, W. R. (2004). Psychoeducation as evidence-based practice: Considerations for practice, research, and policy. *Brief Treatment and Crisis Intervention*, *4*(3), 205–225.

Mahmood, S. S., Wroe, E., Fuller, A., & Leaning, J. (2017). The rohingya people of myanmar: health, human rights, and identity. *The Lancet*, *389*(10081), 1841–1850.

Mensah, F. K. & Kiernan, K. E. (2010). Parents' mental health and children's cognitive and social development. *Social Psychiatry and Psychiatric Epidemiology*, *45*(11), 1023–1035.

Miller, K. E. & Rasmussen, A. (2010). War exposure, daily stressors, and mental health in conflict and post-conflict settings: bridging the divide between trauma-focused and psychosocial frameworks. *Social Science & Medicine*, *70*(1), 7–16.

Motlagh, J. (2018). The Survivors of the Rohingya Genocide. https://www.rollingstone.com/politics/politics-features/rohingya-genocide-myanmar-701354/. Online; Accessed October 9, 2019.

Neugebauer, R., Fisher, P. W., Turner, J. B., Yamabe, S., Sarsfield, J. A., & Stehling-Ariza, T. (2009). Post-traumatic stress reactions among rwandan children and adolescents in the early aftermath of genocide. *International Journal of Epidemiology*, *38*(4), 1033–1045.

Nezlek, J. B., Imbrie, M., & Shean, G. D. (1994). Depression and everyday social interaction. *Journal of Personality and Social Psychology*, *67*(6), 1101.

Ozler, B. (2017). Dealing with attrition in field experiments. https://blogs.worldbank.org/impactevaluations/dealing-attrition-field-experiments. Online; Accessed October 23, 2020.

Patel, V., Rahman, A., Jacob, K., & Hughes, M. (2004). Effect of maternal mental health on infant growth in low income countries: new evidence from south asia. *BMJ*, *328*(7443), 820–823.

Porter, M. & Haslam, N. (2005). Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. *JAMA*, *294*(5), 602–612.

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401.

Rahman, A., Malik, A., Sikander, S., Roberts, C., & Creed, F. (2008). Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural pakistan: a cluster-randomised controlled trial. *The Lancet*, *372*(9642), 902–909.

Rahman, A., Patel, V., Maselko, J., & Kirkwood, B. (2008). The neglected 'm'in mch programmes–why mental health of mothers is important for child nutrition. *Tropical Medicine & International Health*, *13*(4), 579–583.

Raley, S. & Bianchi, S. (2006). Sons, daughters, and family processes: Does gender of children matter? *Annual Review of Sociology*, *32*, 401–421.

Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*, *370*(6522).

Sommer, I., Kien, C., Faustmann, A., Gibson, L., Schneider, M., Krczal, E., Jank, R., & Gartlehner, G. (2018). Prevalence of mental disorders in young refugees and asylum-seekers in european countries. *European Journal of Public Health*, *28*(4), 155–156.

Squires, J. & Bricker, D. (2009). *Ages & Stages Questionnaires: A Parent-Completed Child Monitoring System*. Paul H. Brookes Publishing Co: Baltimore, USA.

Squires, J., Bricker, D., & Twombly, E. (2015). *Ages & stages questionnaires: Social-emotional (A parent-completed child monitoring system for social-emotional behaviors)*. Paul H. Brookes Publishing Co: Baltimore, USA.

Steel, Z., Chey, T., Silove, D., Marnane, C., Bryant, R. A., & Van Ommeren, M. (2009). Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. *JAMA*, *302*(5), 537–549.

Stillman, S., McKenzie, D., & Gibson, J. (2009). Migration and mental health: Evidence from a natural experiment. *Journal of Health Economics*, *28*(3), 677–687.

Summerfield, D. (2000). War and mental health: a brief overview. *BMJ*, *321*(7255), 232–235.

The Lancet (2019). The rohingya people: past, present, and future. *The Lancet*, *394*(2202).

UNHCR (2016). UNHCR Statistcial Yearbook. United Nations High Commissioner for Refugees, Geneva.

UNHCR (2018). Global Trends Forced Displacement in 2018. United Nations High Commissioner for Refugees, Geneva.

UNHCR Camp Profiles (2019). Camp Profiles: Rohingya Refugee Response Bangladesh. United Nations High Commissioner for Refugees, Geneva.

UNHCR Emergencies (2019). Rohingya Emergency. https://www.unhcr.org/uk/rohingya-emergency.html. Online; Accessed October 10, 2019.

UNHCR Population Factsheet (2019). Population Factsheet: Rohingya Refugee Response Bangladesh. United Nations High Commissioner for Refugees, Geneva.

Vlassopoulos, M., Siddique, A., Rahman, T., Pakrashi, D., Islam, A., Ahmed, F., et al. (2021). Improving women's mental health during a pandemic. *Munich Papers in Political Economy no. 02/2021*.

Westfall, P. H. & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons.

WHO (2006). WHO Child Growth Standards: Methods and Development. https://www.who.int/childgrowth/standards/Technical_report.pdf. Online; Accessed December 06, 2019.

WHO (2009). WHO Child Growth Standards: Growth Velocity Based on Weight, Length and Head Circumference: Methods and Development. Geneva, Swtizerland: WHO Department of Nutrition for Health and Development.

World Health Organization (2008). The global burden of disease: 2004 update. World Health Organization, Geneva.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, *134* (2), 557–598.

Main Tables & Figures

Table 1: The home-based psychoeducation program timeline

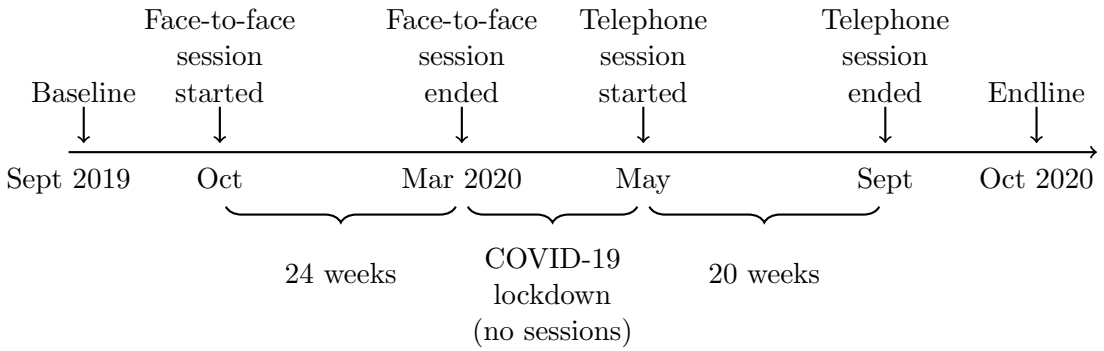
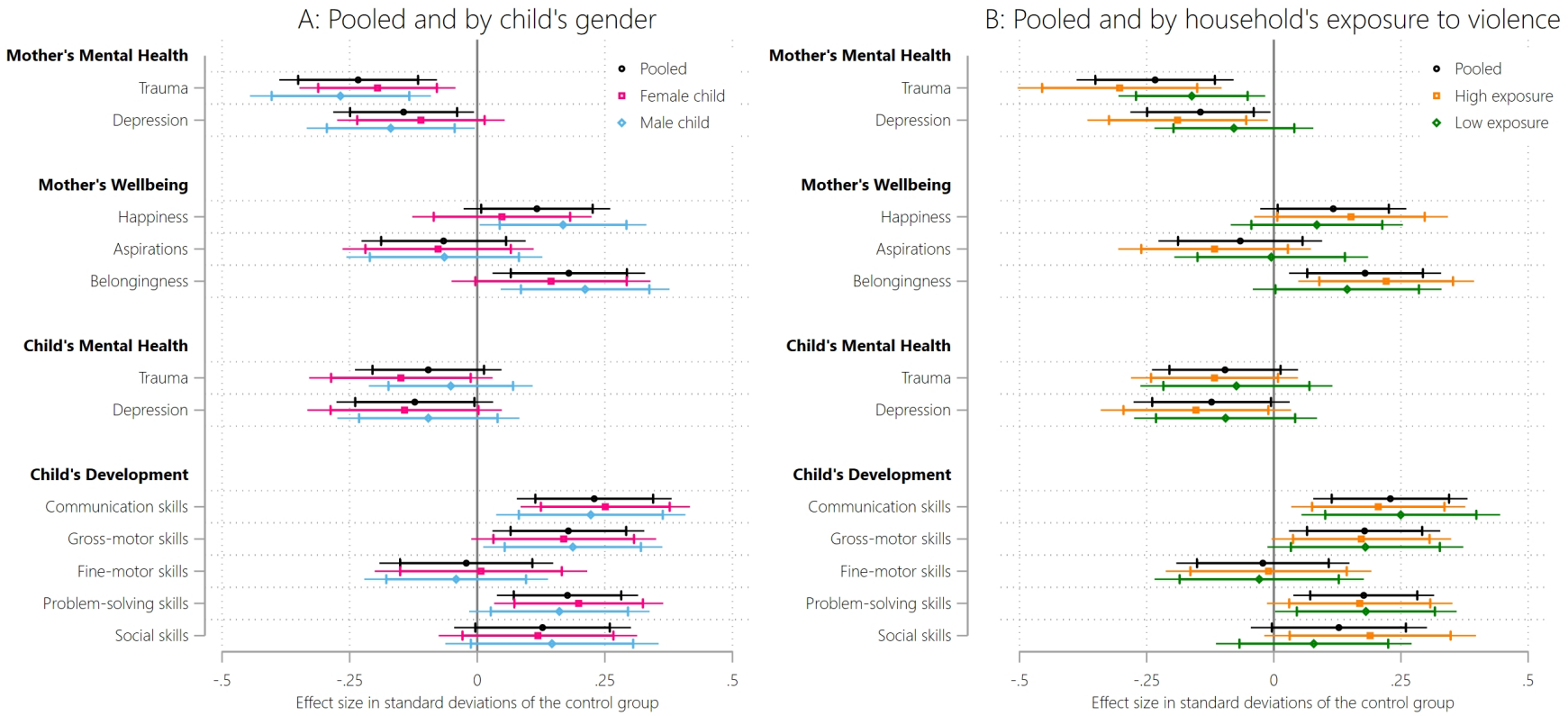


Table 2: Baseline characteristics and balance checks

VARIABLES	**Treatment** (*Std. Dev.*)	N_T	**Control** (*Std. Dev.*)	N_C	**T-test** ***p*-values**	**RI** ***p*-values**
Age of mother	25.70	1,909	25.25	1,586	0.04**	0.03**
	(5.76)		(5.72)			
Mother receives food voucher (=1 if true)	0.51	1,909	0.50	1,586	0.50	0.49
	(0.50)		(0.50)			
Household size	5.30	1,911	5.19	1,586	0.11	0.10
	(2.05)		(1.90)			
Mother employed (=1 if true)	0.02	1,909	0.03	1,586	0.89	0.90
	(0.15)		(0.17)			
Monthly income of mother (=1 if $> 5,000$)	0.41	46	0.49	45	0.65	0.67
	(0.50)		(0.51)			
Husband is alive (=1 if true)	0.97	1,911	0.97	1,586	0.28	0.29
	(0.18)		(0.16)			
Number of children	2.93	1,911	2.90	1,586	0.56	0.55
	(2.00)		(1.89)			
Mother attended school (=1 if true)	0.73	1,910	0.73	1,586	0.83	0.84
	(0.44)		(0.44)			
Months living in the camp	25.00	1,911	26.41	1,586	0.13	0.15
	(8.61)		(18.28)			
Mother is the household head (=1 if true)	0.22	1,911	0.21	1,586	0.53	0.55
	(0.41)		(0.40)			
Mother victim of conflict abuse (=1 if true)	0.87	1,911	0.86	1,586	0.96	0.97
	(0.34)		(0.34)			
Mother victim of camp abuse (=1 if true)	0.16	1,911	0.16	1,586	0.96	0.97
	(0.36)		(0.36)			
Age of child	14.59	1,911	14.23	1,588	0.11	0.11
	(6.44)		(6.50)			
Gender of child	0.50	1,911	0.52	1,588	0.29	0.28
	(0.50)		(0.50)			
Weight of child (in kg)	8.69	1,911	8.57	1,588	0.11	0.12
	(2.19)		(2.16)			
Height of child (in cm)	74.67	1,911	73.98	1,588	0.07*	0.09*
	(9.72)		(9.94)			
Child victim of camp abuse (=1 if true)	0.03	1,911	0.05	1,588	0.38	0.40
	(0.17)		(0.21)			

Note: *Treatment* and *Control* columns show mean of the corresponding variables; all variables with "=1 if true" are dummies and are self explanatory; *Age* is in years; *Household Size* is the number of household members who eat together; *Monthly Income* is a dummy variable that equals 1 if the employed mother earns more than 5,000 Taka per month and 0 if earns less than 5,000 Taka per month (please note that only 91 mothers are employed within the camp); *Months living in the camp* is the number of months the mother have been living in the refugee camp; *Mother victim of conflict abuse* is a dummy variable that equals to 1 if the mother or any household member has experienced at least one type of conflict induced abuse/violence (i.e. either physical, sexual, or verbal abuse, or any harm to the house or the village) and 0 otherwise; *Mother victim of camp abuse* is a dummy variable that equals to 1 if the mother has experienced at least one type of abuse in refugee camps (i.e. either physical, sexual, or verbal abuse); *Child victim of camp abuse* is a dummy variable that equals to 1 if the child has experienced at least one type of abuse in refugee camps (i.e. either physical, sexual, or verbal abuse). T-test p-values are derived from linear regressions, where the dependent variable is from the list above and the independent variable is a dummy that equals 1 if belongs to the treatment group and 0 if not, with camp fixed effects and robust standard errors clustered at the block level; RI p-values are randomization inference p-values (with 1,000 replications) (Young, 2019). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure 1: Treatment effects in standard deviations



Note: This figure shows estimated treatment effects in standard deviations, where the control group has mean 0 and standard deviation 1. Effects reported with 99% and 95% confidence intervals.

Table 3: Treatment effects on mental health and child development

	Treatment effects				
	Without covariates	**With covariates**	**Tr./Dep. at baseline**	**(2)-RI *p*-values**	**(2)-FWER *p*-values**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
A1. Mothers' mental health ‡					
Trauma severity	-0.233***	-0.233***	-0.255***	0.00	0.00
	(0.055)	(0.051)	(0.068)		
Depression severity	-0.146**	-0.144***	-0.288***	0.00	0.02
	(0.057)	(0.054)	(0.095)		
Composite mental health index	-0.223***	-0.223***	-0.276***	0.00	0.00
	(0.059)	(0.054)	(0.072)		
A2. Mothers' well-being					
Happiness	0.108*	0.117**	-	0.04	0.04
	(0.057)	(0.056)			
Aspirations	-0.068	-0.066	-	0.32	0.69
	(0.062)	(0.062)			
Belongingness	0.180***	0.179***	-	0.00	0.00
	(0.058)	(0.057)			
Composite SWB index	0.116**	0.119**	-	0.04	0.02
	(0.057)	(0.055)			
B1. Children's mental health ‡					
Trauma severity	-0.117**	-0.096*	-0.127*	0.08	0.02
	(0.057)	(0.055)	(0.074)		
Depression severity	-0.128**	-0.122**	-0.239**	0.03	0.02
	(0.061)	(0.059)	(0.098)		
Composite mental health index	-0.139**	-0.123**	-0.153**	0.03	0.01
	(0.061)	(0.059)	(0.073)		
B2. Children's development					
Communication skills	0.251***	0.229***	-	0.00	0.00
	(0.061)	(0.059)			
Gross-motor skills	0.197***	0.179***	-	0.00	0.00
	(0.061)	(0.058)			
Fine-motor skills	0.006	-0.021	-	0.76	0.89
	(0.071)	(0.066)			
Problem-solving skills	0.195***	0.177***	-	0.00	0.00
	(0.058)	(0.055)			
Social skills	0.125*	0.128*	-	0.05	0.01
	(0.067)	(0.067)			
Composite child development index	0.203***	0.182***	-	0.00	0.00
	(0.072)	(0.069)			
Observations	2,845	2,840	$1{,}240^T/508^D$	-	-

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1): treatment effect estimated without any baseline covariates. Column (2): treatment effect estimated with all baseline covariates (as in equation 1). Column (3): treatment effect only on mothers that were found to be traumatized ($N = 1,240$)/depressed ($N = 508$) at the baseline, with all covariates. All outcomes are standardized indices, such that the control group has mean 0 and standard deviation 1. The composite indices aggregate the individual outcome indices under each panel. For mental health outcomes (under A1 and B1), lower values correspond to improvement in mental health. For other outcomes (under A2 and B2), higher values correspond to more favorable outcomes. Covariates include baseline measures of age (mother's and child's), whether mother attend school, household size, monthly household spending, months lived in the camp, whether mother receives monthly food voucher, whether child's father is alive, any family member stranded in Myanmar, gender of the child, number of children, household victimization (based on household's experience during conflict in Myanmar), mothers' camp-victimization (based on abuse in the camp), and children's camp-victimization (based on abuse in the camp). Standard errors, clustered at the block level (251 clusters), are in parentheses. Column (4) reports RI *p*-values for the full model (column 2), which are randomization inference *p*-values (with 1,000 replications) (Young, 2019). Column (5) reports FWER *p*-values for the full model (column 2), which are the List-Shaikh-Xu familywise error rate adjusted p-values (with 3,000 replications) based on 12 tests (List et al., 2019).

Table 4: Mentally unwell in treatment arm versus mentally healthy in control arm: Are the treated catching up?

	X: Trauma		Y: Depression	
	Without covariates	With covariates	Without covariates	With covariates
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
A1. Mother's mental health ‡				
Trauma severity	-0.190***	-0.200***	-0.131	-0.136
	(0.067)	(0.073)	(0.100)	(0.122)
Depression severity	-0.093	-0.106	0.010	0.041
	(0.063)	(0.067)	(0.085)	(0.106)
A2. Mother's well-being				
Happiness	0.107	0.117	0.195**	0.243**
	(0.073)	(0.076)	(0.094)	(0.107)
Aspirations	-0.069	-0.075	-0.026	-0.049
	(0.079)	(0.078)	(0.096)	(0.102)
Belongingness	0.204***	0.207***	0.308***	0.351***
	(0.074)	(0.072)	(0.096)	(0.093)
B1. Child's mental health ‡				
Trauma severity	-0.142**	-0.122*	-0.118	-0.110
	(0.072)	(0.071)	(0.087)	(0.091)
Depression severity	-0.161**	-0.162**	-0.233***	-0.297***
	(0.080)	(0.082)	(0.088)	(0.091)
B2. Child's development				
Communication skills	0.210***	0.158**	0.285***	0.277***
	(0.078)	(0.076)	(0.108)	(0.104)
Gross-motor skills	0.216***	0.190**	0.285***	0.327***
	(0.078)	(0.079)	(0.105)	(0.108)
Fine-motor skills	0.092	0.043	0.162	0.161
	(0.091)	(0.085)	(0.126)	(0.114)
Problem-solving skills	0.258***	0.211***	0.282***	0.230**
	(0.073)	(0.071)	(0.098)	(0.095)
Social skills	0.096	0.100	0.216*	0.281**
	(0.086)	(0.087)	(0.116)	(0.115)
Observations	1,405	1,405	852	852

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Vertical panel X (trauma) includes participants from the treatment arm that were traumatized at baseline (or the mentally unwell) and participants from the control arm that did not have any trauma symptoms at baseline (or the mentally well). Similarly, vertical panel Y (depression) includes participants from the treatment arm that were depressed at baseline (or the mentally unwell) and participants from the control arm that did not have any depressive symptoms at baseline (or the mentally well). Columns (1) and (3): treatment effects estimated without any baseline covariates. Columns (2) and (4): treatment effect estimated with all baseline covariates (as in equation 3). Covariates are listed under Table 3. For outcomes with ‡, negative coefficients imply more favorable outcomes.

Table 5: Treatment effects on child malnutrition

		Treatment effects				
	Control mean	**Without covariates**	**With all covariates**	**Girl child**	**Boy child**	**Diff (5)-(4)**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Underweight						
Weight-for-age z-score (WAZ)	-3.17	0.630***	0.641***	0.764***	0.551***	-0.115
	[2.34]	(0.093)	(0.092)	(0.120)	(0.125)	(0.162)
Weight (in kg)	8.75	0.700***	0.635***	0.778***	0.519***	-0.172
	[3.23]	(0.130)	(0.123)	(0.158)	(0.168)	(0.220)
Underweight (=1 if $WAZ < -2$)	0.73	-0.070***	-0.071***	-0.107***	-0.046*	0.042
	[0.44]	(0.019)	(0.019)	(0.026)	(0.026)	(0.035)
Severe underweight (=1 if $WAZ < -3$)	0.60	-0.150***	-0.158***	-0.185***	-0.134***	0.033
	[0.49]	(0.020)	(0.020)	(0.028)	(0.027)	(0.037)
B. Stunting						
Height-for-age z-score (HAZ)	-2.66	0.647***	0.515***	0.645***	0.417**	0.015
	[3.77]	(0.153)	(0.139)	(0.192)	(0.193)	(0.256)
Height (in cm)	80.5	2.366***	1.576***	2.090***	1.156*	-0.185
	[13.91]	(0.625)	(0.454)	(0.640)	(0.628)	(0.855)
Stunting (=1 if $HAZ < -2$)	0.69	-0.081***	-0.070***	-0.081***	-0.063**	-0.015
	[0.46]	(0.019)	(0.018)	(0.028)	(0.026)	(0.038)
Severe stunting (=1 if $HAZ < -3$)	0.60	-0.143***	-0.130***	-0.132***	-0.132***	-0.038
	[0.49]	(0.020)	(0.019)	(0.028)	(0.028)	(0.039)
C. Wasting						
Weight-for-height z-score (WHZ)	-2.43	0.423***	0.508***	0.464***	0.566***	0.019
	[3.23]	(0.135)	(0.125)	(0.156)	(0.182)	(0.213)
Wasting (=1 if $WHZ < -2$)	0.57	-0.059***	-0.081***	-0.056**	-0.102***	-0.029
	[0.49]	(0.022)	(0.018)	(0.028)	(0.025)	(0.035)
Severe wasting (=1 if $WHZ < -3$)	0.43	-0.108***	-0.128***	-0.089***	-0.159***	-0.053*
	[0.50]	(0.022)	(0.018)	(0.024)	(0.026)	(0.032)
Observations	1,166	2,845	2,840	1,400	1,440	2,840

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1): control group average at endline with standard deviations in brackets; Column (2): treatment effect estimated without any baseline covariates. Column (3): treatment effect estimated with all baseline covariates (as in equation 1). Column (4): treatment effect on girl child. Column (5): treatment effect on boy child. Column (6): difference between column (4) and (5), which is the coefficient on the interaction between treatment dummy and child's gender dummy. Average age of child at endline was 27 months. For z-scores, higher values correspond to more favorable outcomes. For indicators, lower values correspond to more favorable outcomes. Standard errors, clustered at the block level (251 clusters), are in parentheses. Column (4) reports RI p-values for the full model (column 2), which are randomization inference p-values (with 1,000 replications) (Young, 2019). Column (5) reports FWER p-values for the full model (column 2), which are the List-Shaikh-Xu familywise error rate adjusted p-values (with 3,000 replications) based on 12 tests (List et al., 2019).

Table 6: Treatment effect on transmission of mental health, pooled and by gender of child

	Treatment effects			
	Without covariates	**With covariates**	**(2)-RI *p*-values**	**(2)-FWER *p*-values**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
A1. Trauma, pooled				
Difference	-0.188***	-0.177***	0.00	0.00
	(0.056)	(0.054)		
A2. Trauma, by child's gender				
Difference, if girl	-0.157**	-0.147**	0.03	-
	(0.066)	(0.066)		
Difference, if boy	-0.221***	-0.216***	0.00	-
	(0.070)	(0.069)		
B1. Depression severity, pooled				
Difference	-0.157**	-0.155**	0.03	0.00
	(0.072)	(0.069)		
B2. Depression severity, by child's gender				
Difference, if girl	-0.167*	-0.167**	0.05	-
	(0.086)	(0.084)		
Difference, if boy	-0.141*	-0.134*	0.09	-
	(0.081)	(0.079)		
Observations	2,803	2,798	-	-

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1): treatment effect estimated without any baseline covariates. Column (2): treatment effect estimated with all baseline covariates (as in equation 3). Covariates are listed under Table 3. Dependent variables are absolute differences in mental health indices (trauma under A1 and A2, and depression under B1 and B2) between mothers and children. That is, $Difference = |Mother - Child|$'s mental health index. Standard errors, clustered at the block level (251 clusters), are in parentheses. Column (3) reports RI *p*-values for the full model (column 2), which are randomization inference *p*-values (with 1,000 replications) (Young, 2019). Column (4) reports FWER *p*-values for the full model (column 2), which are the List-Shaikh-Xu (LSX) familywise error rate adjusted p-values (with 3,000 replications) (List et al., 2019). Since LSX does not allow covariates or conditions (e.g., by gender of the child), we report FWER-adjusted *p*-values only for the pooled observations. Observations with girl child is 1,387 in column (1) and 1,382 in column (2). Observations with boy child is 1,416 in both columns (1) and (2). Correlation of mother-child mental health is reported in Table A5 in Appendix A.

Table 7: Social desirability bias check

	Mother outcomes					Child outcomes						
VARIABLES	**Trauma (1)**	**Dep. (2)**	**Happ. (3)**	**Aspr. (4)**	**Belong. (5)**	**Trauma (6)**	**Dep. (7)**	**Comm. (8)**	**Gross. (9)**	**Fine. (10)**	**Prob. (11)**	**Social. (12)**
Treatment	-0.229***	-0.123**	0.114**	-0.060	0.200***	-0.068	-0.134**	0.208***	0.158**	-0.035	0.144**	0.144*
	(0.055)	(0.061)	(0.057)	(0.067)	(0.067)	(0.061)	(0.067)	(0.070)	(0.070)	(0.071)	(0.063)	(0.080)
High SDB	0.083	0.042	-0.024	0.014	-0.009	0.100	0.014	0.025	-0.041	-0.019	-0.039	0.063
	(0.050)	(0.058)	(0.054)	(0.057)	(0.062)	(0.061)	(0.063)	(0.056)	(0.063)	(0.057)	(0.058)	(0.058)
Treatment×High SDB	-0.008	-0.050	0.005	-0.013	-0.048	-0.065	0.027	0.051	0.049	0.032	0.076	-0.036
	(0.061)	(0.068)	(0.068)	(0.075)	(0.077)	(0.077)	(0.076)	(0.069)	(0.078)	(0.073)	(0.073)	(0.076)
Observations	2,798	2,798	2,798	2,798	2,798	2,798	2,798	2,840	2,840	2,840	2,840	2,840
R-squared	0.040	0.026	0.028	0.063	0.062	0.032	0.017	0.081	0.054	0.093	0.081	0.026

Robust standard errors clustered at the block level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All outcomes are standardized indices such that the control group has mean zero and SD one. Outcomes in columns 1-5 are of mothers: (1) trauma, (2) depression, (3) happiness, (4) future aspirations, and (5) belongingness. Outcomes in columns 6-12 are of children: (6) trauma, (7) depression, (8) communication skills, (9) gross-motor skills, (10) fine-motor skills, (11) problem-solving skills, and (12) social skills. Treatment is a dummy that equals to 1 if respondents are in the treatment arm and 0 otherwise. High SDB is a dummy that equals to 1 if the social desirability bias (SDB) score is above 8 (which is the median value) and 0 if below. All specifications include the usual set of controls and camp fixed effects as in Table 3.

Table 8: Potential mechanisms

		Treatment effects			
	Control mean	Pooled	Girl child	Boy child	Diff (4)-(3)
Intermediate outcomes	(1)	(2)	(3)	(4)	(5)
A. Mental health of mothers					
Doctor visits (0-4)	1.88	0.004	0.014	-0.011	-0.027
	[0.79]	(0.034)	(0.045)	(0.045)	(0.059)
Disagreements/arguments with spouse (0-4)	1.04	-0.054	-0.070	-0.038	0.022
	[0.90]	(0.034)	(0.053)	(0.045)	(0.068)
Seek help for household chores (0-4)	1.05	-0.016	0.004	-0.041	-0.030
	[0.95]	(0.039)	(0.058)	(0.055)	(0.078)
Communication during lockdown (0-4)	1.93	-0.011	0.011	-0.023	0.005
	[0.78]	(0.029)	(0.041)	(0.043)	(0.055)
B. Children's development					
Mother's time input per day (0-24)	9.15	1.498***	1.915***	1.113***	-0.684
	[5.83]	(0.244)	(0.324)	(0.331)	(0.436)
Father's time input per day (0-24)	5.14	0.066	-0.053	0.144	0.215
	[3.01]	(0.114)	(0.168)	(0.160)	(0.226)
Age stopped breastfeeding	20.83	0.161	-0.161	0.414*	0.653*
	[5.04]	(0.173)	(0.267)	(0.250)	(0.361)
Times feeding child per day	3.97	0.011	0.041	-0.017	-0.074
	[1.47]	(0.057)	(0.080)	(0.074)	(0.104)
Negative parenting (0-4)	0.67	-0.022*	-0.027	-0.016	0.004
	[0.33]	(0.011)	(0.017)	(0.014)	(0.022)
Ask others to babysit (0-4)	0.87	0.011	0.035	-0.007	-0.060
	[0.94]	(0.038)	(0.058)	(0.052)	(0.071)
Prevalence of indoor smoking (0-4)	0.32	0.036	0.067	0.006	-0.028
	[0.76]	(0.030)	(0.044)	(0.041)	(0.059)
Let child walk/play barefoot (0-4)	0.65	-0.069**	-0.029	-0.117***	-0.056
	[0.83]	(0.032)	(0.046)	(0.042)	(0.059)
Observations	1,166	2,840	1,400	1,440	2,840

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1): control group average at endline with standard deviations in brackets; Column (2): treatment effect estimated without any baseline covariates. Column (3): treatment effect estimated with all baseline covariates (as in equation 1). Standard errors, clustered at the block level (251 clusters), are in parentheses. Column (4) reports RI p-values for the full model (column 2), which are randomization inference p-values (with 1,000 replications) (Young, 2019). Column (5) reports FWER p-values for the full model (column 2), which are the List-Shaikh-Xu familywise error rate adjusted p-values (with 3,000 replications) based on 12 tests (List et al., 2019).

Table 9: Causal mediation analysis

	Trauma of children		Depression of children	
	Means	**95% CI**	**Means**	**95% CI**
Effects	**(1)**	**(2)**	**(3)**	**(4)**
	Panel A: Mediator - Trauma of Mothers			
Direct	-0.02	[-0.12, 0.07]	-0.06	[-0.16, 0.04]
ACME	-0.06***	[-0.08, -0.03]	-0.04***	[-0.07, -0.02]
	Panel B: Mediator - Depression of Mothers			
Direct	-0.09	[-0.20, 0.01]	-0.06	[-0.16, 0.04]
ACME	-0.02**	[-0.03, -0.00]	-0.02**	[-0.04, -0.00]
	Panel C: Mediator - Happiness of Mothers			
Direct	-0.07	[-0.18, 0.03]	-0.09	[-0.21, 0.01]
ACME	-0.01	[-0.02, 0.00]	-0.01	[-0.03, 0.00]
	Panel D: Mediator - Aspirations of Mothers			
Direct	-0.08	[-0.19, 0.02]	-0.10	[-0.22, 0.01]
ACME	-0.00	[-0.01, 0.00]	-0.00	[-0.01, 0.00]
	Panel E: Mediator - Belongingness of Mothers			
Direct	-0.05	[-0.16, 0.05]	-0.09	[-0.21, 0.02]
ACME	-0.02**	[-0.04, -0.01]	-0.02**	[-0.03, -0.00]

*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variables are trauma and depression indices of children (same as in Table 3). Columns 1 and 3 reports the effect sizes estimated using OLS. Coefficients of 'Direct' reports the direct treatment effect. ACME reports the average causal mediation effect (Imai et al., 2010). Columns with '95% CI' reports the 95% confidence interval. All regressions control for the usual individual and household characteristics as in Table 3. Also, standard errors are clustered at the unit of randomization (i.e., blocks).

Table 10: Treatment effects: Inverse Probability Weighting & Lee bounds

	Treatment effects		Lee (2009) bounds	
	Unadjusted	IPW	95% CI	95% CI (Tightened)
Dependent variables	(1)	(2)	(3)	(4)
A1. Mothers' mental health				
Trauma	-0.233***	-0.234***	[-0.528, -0.100]	[-0.519, -0.101]
	(0.051)	(0.049)		
Depression	-0.144***	-0.144***	[-0.381, -0.047]	[-0.383, -0.059]
	(0.054)	(0.052)		
A2. Mothers'wellbeing				
Happiness	0.117**	0.124**	[-0.055, 0.596]	[-0.048, 0.591]
	(0.056)	(0.054)		
Aspirations	-0.066	-0.073	[-0.360, 0.318]	[-0.246, 0.294]
	(0.062)	(0.061)		
Belongingness	0.179***	0.190***	[0.006 0.562]	[0.017, 0.568]
	(0.057)	(0.055)		
B1. Children's mental health				
Trauma	-0.096*	-0.094*	[-0.439, 0.039]	[-0.388, 0.035]
	(0.055)	(0.054)		
Depression	-0.122**	0.117**	[-0.391, 0.003]	[-0.366, -0.003]
	(0.059)	(0.057)		
B2. Children's development				
Communication skills	0.229***	0.232***	[0.070, 0.689]	[0.054, 0.645]
	(0.059)	(0.058)		
Gross-motor skills	0.179***	0.189***	[0.106, 0.555]	[0.103, 0.530]
	(0.058)	(0.056)		
Fine-motor skills	-0.021	-0.017	[-0.356, 0.364]	[-0.340, 0.371]
	(0.066)	(0.064)		
Problem-solving skills	0.177***	0.172***	[-0.104, 0.564]	[-0.089. 0.527]
	(0.055)	(0.054)		
Social skills	0.128*	0.148**	[-0.204, 0.480]	[-0.245, 0.490]
	(0.067)	(0.066)		

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1) reports unadjusted/unweighted treatment effects, same as in Table 3. Column (2) reports the Inverse Probability Weight (IPW) adjusted treatment effects. Columns (3)-(4) report 95 percent confidence (CI) intervals on the treatment effects using Lee (2009) bounds, where Column (4) reports CI with tighter bounds using two covariates that explains attrition across treatment and control groups (see Table A8 in Appendix A).

Table 11: Heterogeneous treatment effects, by gender and violence exposure

		by child's gender			by violence exposure		
	Pooled	**Girl**	**Boy**	**Difference (β_3)**	**High**	**Low**	**Difference (β_3)**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
A1. Mothers' mental health							
Trauma	-0.233***	-0.195***	-0.268***	-0.068	-0.303***	-0.161***	-0.157*
	(0.051)	(0.061)	(0.064)	(0.070)	(0.072)	(0.056)	(0.082)
Depression	-0.144***	-0.110*	-0.170***	-0.056	-0.189***	-0.079	-0.130*
	(0.054)	(0.063)	(0.063)	(0.062)	(0.068)	(0.060)	(0.070)
A2. Mothers' wellbeing							
Happiness	0.117**	0.048	0.168***	0.102	0.152**	0.085	0.064
	(0.056)	(0.068)	(0.063)	(0.065)	(0.073)	(0.065)	(0.082)
Aspirations	-0.066	-0.077	-0.064	0.006	-0.116	-0.005	-0.145*
	(0.062)	(0.072)	(0.074)	(0.074)	(0.073)	(0.074)	(0.080)
Belongingness	0.179***	0.145*	0.211***	0.084	0.221***	0.144**	0.058
	(0.057)	(0.075)	(0.062)	(0.073)	(0.065)	(0.072)	(0.082)
B1. Children's mental health							
Trauma	-0.096*	-0.150**	-0.052	0.065	-0.117*	-0.074	-0.010
	(0.055)	(0.069)	(0.062)	(0.069)	(0.063)	(0.073)	(0.079)
Depression	-0.122**	-0.142*	-0.096	0.006	-0.153**	-0.095	-0.029
	(0.059)	(0.074)	(0.069)	(0.073)	(0.072)	(0.070)	(0.079)
B2. Children's development							
Communication skills	0.229***	0.251***	0.222***	-0.007	0.205***	0.250***	-0.083
	(0.059)	(0.070)	(0.071)	(0.074)	(0.066)	(0.076)	(0.081)
Gross-motor skills	0.179***	0.169**	0.187***	0.015	0.172**	0.180**	-0.048
	(0.058)	(0.070)	(0.068)	(0.075)	(0.068)	(0.074)	(0.083)
Fine-motor skills	-0.021	0.007	-0.041	-0.063	-0.010	-0.029	-0.016
	(0.066)	(0.081)	(0.070)	(0.075)	(0.078)	(0.080)	(0.084)
Problem-solving skills	0.177***	0.199***	0.161**	-0.010	0.169**	0.181***	-0.041
	(0.055)	(0.062)	(0.068)	(0.069)	(0.071)	(0.069)	(0.084)
Social skills	0.128*	0.119	0.146*	0.011	0.189**	0.079	0.044
	(0.067)	(0.075)	(0.081)	(0.077)	(0.080)	(0.074)	(0.080)
Observations	2,798	1,382	1,416	2,798	1,457	1,341	2,798

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. Columns 2-3 and 5-6 report estimates from split samples. For instance, estimates in column 2 (3) are derived from the sample with only female (male) children. Columns 4 and 7 report the coefficients on the interaction term from equation 5.

Forced Displacement, Mental Health, and Child Development: Evidence from the Rohingya Refugees

Online Appendix

By Asad Islam[1], Tanvir Ahmed Mozumder[2], Tabassum Rahman[3], Tanvir Shatil[4], Abu Siddique[5]

List of Figures

List of Tables

[1]Centre for Development Economics & Sustainability (CDES) and Department of Economics, Monash University, Australia; J-PAL. Email: asadul.islam@monash.edu

[2]BRAC Institute of Governance & Development, BRAC University, Bangladesh. Email: tanvir.mozumder@bracu.ac.bd

[3]School of Medicine & Public Health, University of Newcastle, Australia. Email: tabassum.rahman@uon.edu.au

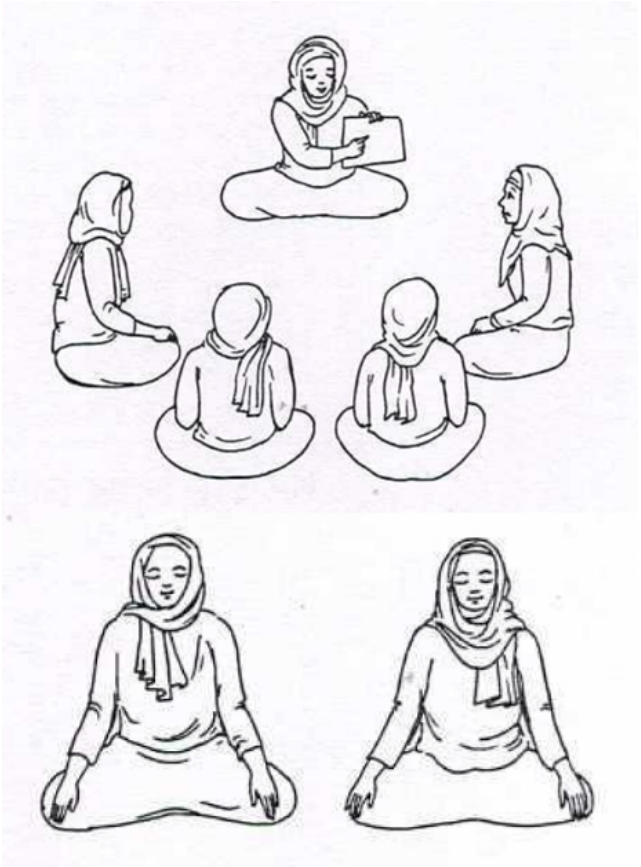
[4]BRAC Institute of Governance & Development, BRAC University, Bangladesh. Email: tanvir.shatil@bracu.ac.bd

[5]Department of Political Economy, King's College London, UK; CDES. Email: abu.b.siddique@kcl.ac.uk

A Appendix: Additional Tables and Figures

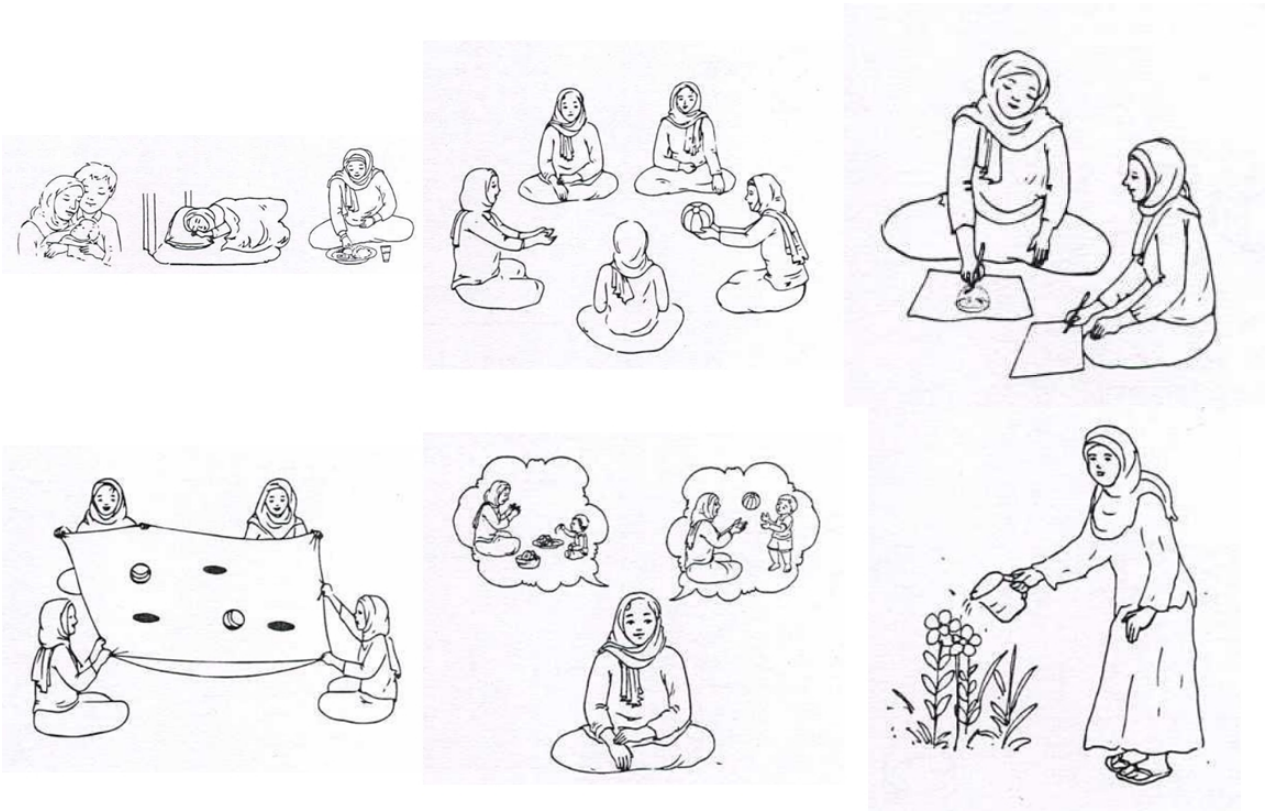
A.1 Figures

Figure A1: Psychoeducation session: step 1



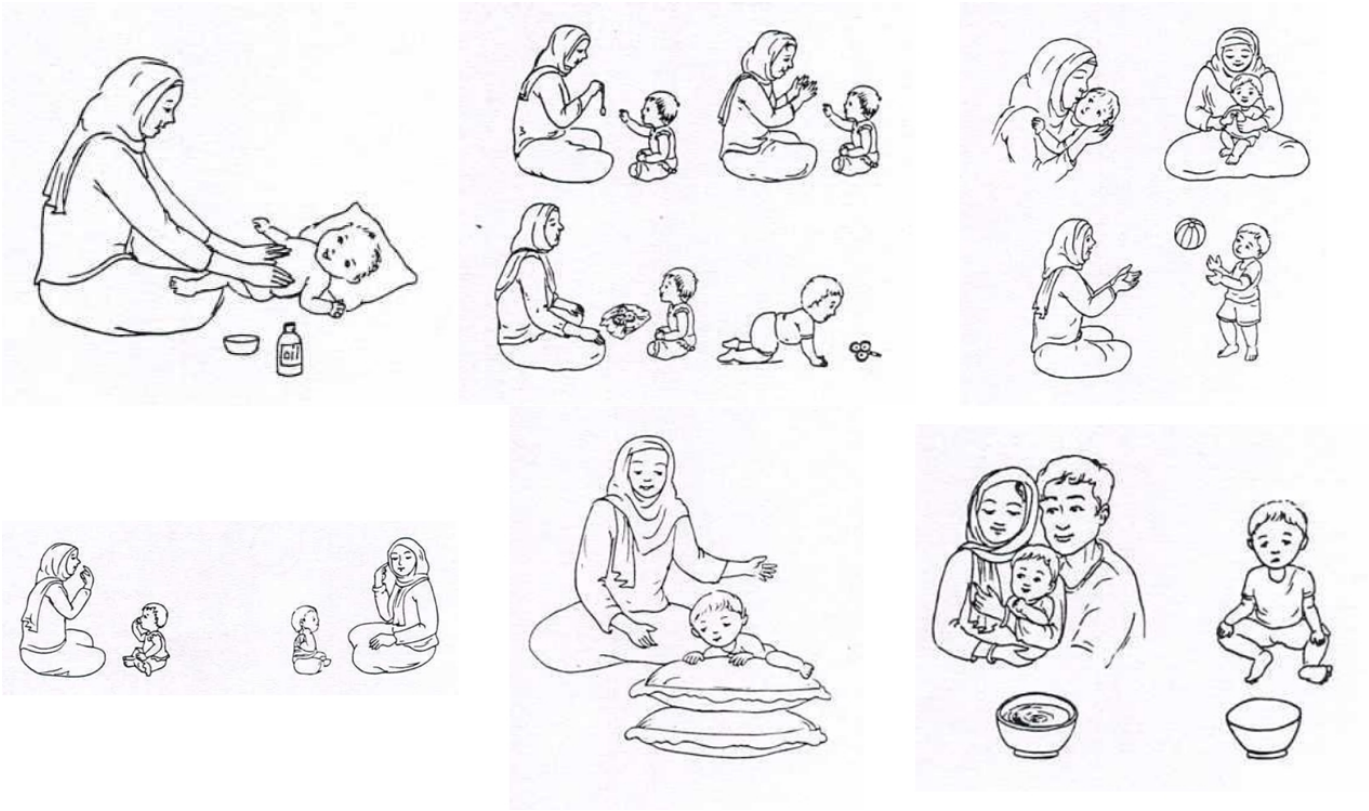
Note: In step 1, mothers greet each other, do breathing exercises, and provide feedback on the homework.

Figure A2: Psychoeducation session: step 2



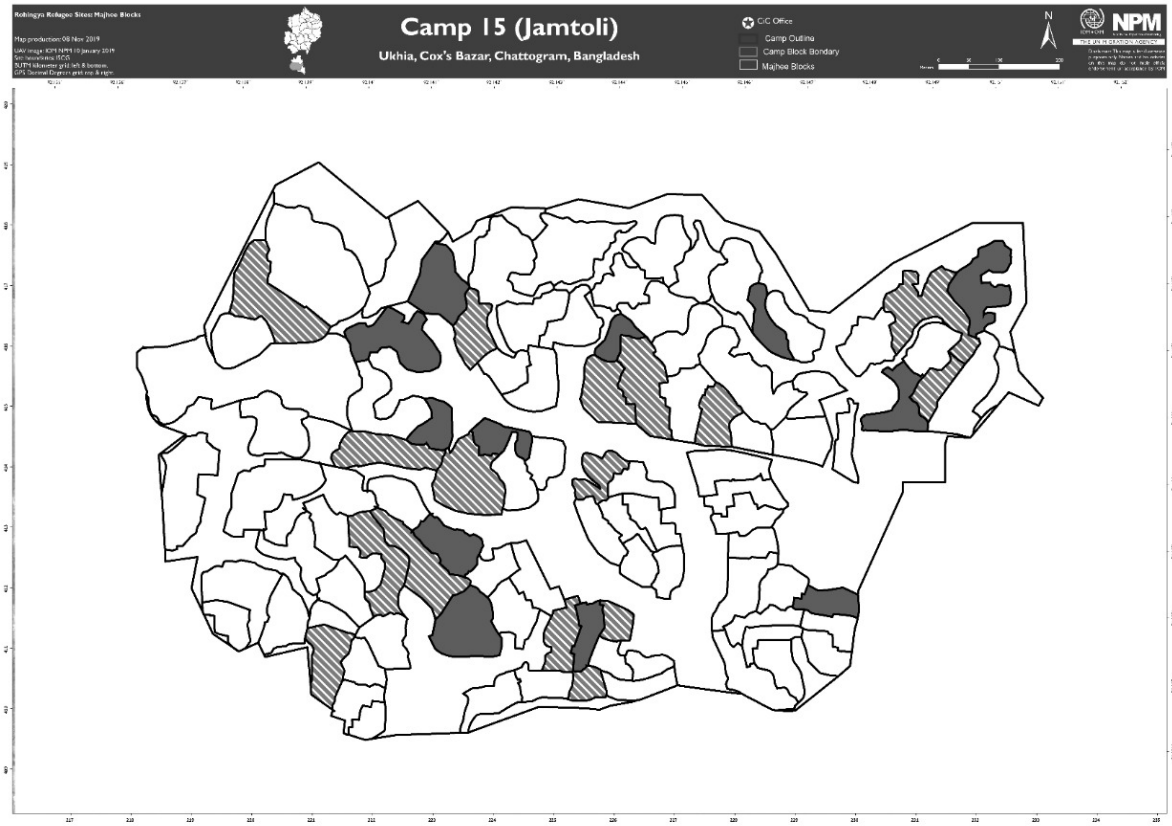
Note: In step 2, mothers engage in discussions on their personal well-being and play games with other mothers.

Figure A3: Psychoeducation session: step 3



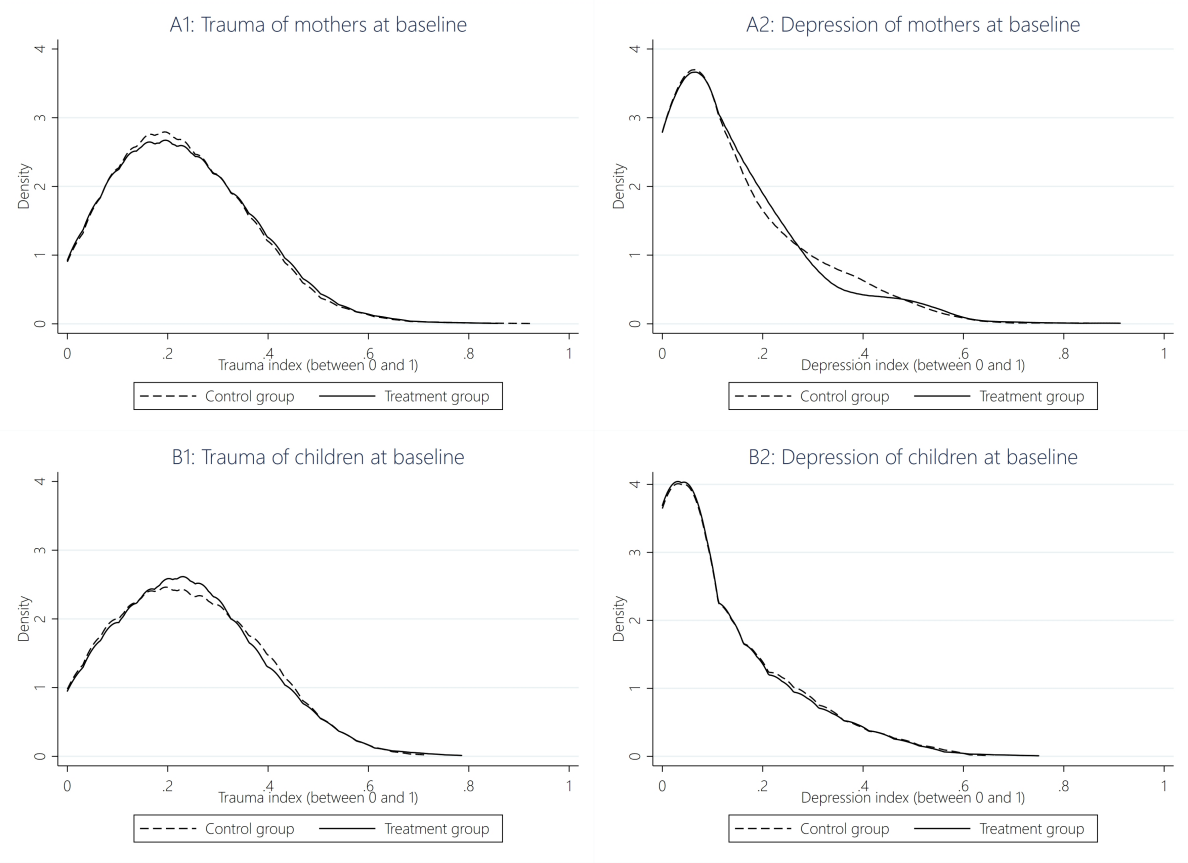
Note: In step 3, mothers receive advice on childcare and engage in play activities with their children during the session.

Figure A4: Map of a Rohingya camp



Note: This is a map of a camp showing the treatment and control blocks. Shaded areas are treatment and dark grey areas are control blocks.

Figure A5: Mental health of mothers and children at baseline



Note: This figure shows the distribution of mental health of mothers (A1 and A2) and children (B1 and B2) at the baseline (estimated from kernel density estimation). Trauma and depression indices are averages of responses to trauma and depression questions, where higher values correspond to more severe mental health conditions. For details on how these two indices are constructed, see the note under Table **??**.

A.2 Tables

Table A1: Baseline outcomes and balance checks

VARIABLES	**Treatment** (*Std. Dev.*)	N_T	**Control** (*Std. Dev.*)	N_C	**T-test** ***p*-values**	**RI** ***p*-values**
	Panel A: Mother outcomes					
Traumatized (=1 if true)	0.45	1,911	0.44	1,586	0.69	0.70
	(0.50)		(0.50)			
Depressed (=1 if true)	0.17	1,911	0.20	1,586	0.87	0.88
	(0.37)		(0.40)			
Happiness index ($0 \leq index \leq 1$)	0.77	1,911	0.78	1,586	0.52	0.53
	(0.17)		(0.17)			
Aspiration index ($0 \leq index \leq 1$)	0.61	1,911	0.62	1,586	0.21	0.21
	(0.11)		(0.11)			
Belongingness index ($0 \leq index \leq 1$)	0.79	1,911	0.79	1,586	0.23	0.20
	(0.15)		(0.16)			
	Panel B: Child outcomes					
Traumatized (=1 if true)	0.49	1,911	0.48	1,588	0.57	0.57
	(0.50)		(0.50)			
Depressed (=1 if true)	0.17	1,911	0.18	1,588	0.97	0.97
	(0.37)		(0.38)			
Communication skills index ($0 \leq index \leq 1$)	0.56	1,911	0.56	1,588	0.75	0.73
	(0.30)		(0.31)			
Gross-motor skills index ($0 \leq index \leq 1$)	0.63	1,911	0.63	1,588	0.86	0.85
	(0.32)		(0.33)			
Fine-motor skills index ($0 \leq index \leq 1$)	0.50	1,911	0.48	1,588	0.44	0.43
	(0.31)		(0.31)			
Problem-solving skills index ($0 \leq index \leq 1$)	0.47	1,911	0.48	1,588	0.60	0.96
	(0.31)		(0.33)			
Social skills index ($0 \leq index \leq 1$)	0.58	1,911	0.59	1,588	0.72	0.64
	(0.28)		(0.29)			
Underweight for age (=1 if true)	0.23	1,911	0.24	1,588	0.52	0.49
	(0.42)		(0.42)			
Severely underweight for age (=1 if true)	0.08	1,911	0.08	1,588	0.97	0.96
	(0.27)		(0.26)			
Stunted for age (=1 if true)	0.27	1,911	0.27	1,588	0.56	0.58
	(0.44)		(0.45)			
Severely stunted for age (=1 if true)	0.13	1,911	0.12	1,588	0.80	0.80
	(0.33)		(0.32)			
Wasted for age (=1 if true)	0.	1,911	0.	1,588	0.52	0.49
	(0.42)		(0.42)			
Severely wasted for age (=1 if true)	0.	1,911	0.	1,588	0.97	0.96
	(0.27)		(0.26)			

Note: *Treatment* and *Control* columns show mean of the corresponding variables. Variables that are indices are averages of responses to survey questions associated with the outcomes, such that the value of each variable is between 0 and 1. For instance, *Communication skills* is measured using 6 questions and each question is answered as either 'yes' (=1) or 'no' (=0). So, the *Communication skills* variable under Panel B simply adds up responses and divides the total by 6 (the highest total score). All index variables have been generated in this way. Therefore, these variables simply show the averages. All variables with "=1 if true" are dummies and are self explanatory; T-test p-values are derived from linear regressions, where the dependent variable is from the list above and the independent variable is a dummy that equals to 1 if belongs to the treatment group and 0 if belongs to the control group with camp fixed effects and robust standard errors clustered at the block level; RI p-values are randomization inference p-values (with 1,000 replications) (Young, 2019). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A2: Impact on child development at different ages

	0-2	2-4	4-6	6-8	8-10	10-12	12+
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Communication skills	0.053	-0.270	0.418	-0.035	0.410	0.205	0.588***
	(0.997)	(0.517)	(0.311)	(0.395)	(0.365)	(0.296)	(0.111)
Control mean	3.028	2.622	2.654	2.822	2.495	3.380	3.219
	[1.823]	[1.512]	[1.745]	[1.994]	[1.882]	[2.050]	[1.772]
Gross-motor skills	-0.560	-1.169**	0.321	0.118	0.438	-0.023	0.460***
	(0.820)	(0.541)	(0.339)	(0.392)	(0.346)	(0.257)	(0.122)
Control mean	2.971	3.800	3.740	3.839	3.114	3.711	3.908
	[2.407]	[2.160]	[1.880]	[1.952]	[1.992]	[1.846]	[1.881]
Fine-motor skills	0.385	-1.025*	0.264	0.146	0.210	-0.453*	-0.115
	(0.759)	(0.527)	(0.346)	(0.324)	(0.384)	(0.242)	(0.162)
Control mean	1.800	2.756	2.442	2.415	2.467	2.818	3.343
	[1.779]	[2.123]	[1.717]	[1.799]	[1.932]	[1.713]	[2.139]
Problem-solving skills	-0.129	-0.388	0.569	-0.137	0.793**	0.237	0.499***
	(0.752)	(0.714)	(0.368)	(0.368)	(0.393)	(0.269)	(0.146)
Control mean	2.314	2.778	2.433	2.822	2.848	3.347	3.161
	[2.180]	[1.857]	[2.028]	[1.929]	[2.156]	[2.283]	[2.218]
Social skills	-0.011	-0.129	0.462	-0.121	0.121	-0.279	0.366***
	(0.785)	(0.522)	(0.361)	(0.342)	(0.321)	(0.267)	(0.129)
Control mean	3.114	4.044	3.856	3.568	3.200	3.397	3.359
	[2.311]	[1.705]	[1.943]	[1.712]	[1.751]	[1.846]	[1.711]
Composite child development index	0.073	-3.126	1.993	-0.107	2.029	-0.230	1.796***
	(3.856)	(2.510)	(1.390)	(1.553)	(1.449)	(0.994)	(0.549)
Control mean	13.229	16.000	15.125	15.466	14.124	16.653	16.991
	[8.809]	[7.793]	[7.532]	[7.802]	[7.643]	[6.705]	[7.176]
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Camp FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52	92	182	180	182	289	1,558

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported (with standard errors in parentheses). Dependent variables are continuous (values between 0 and 6). Each column reports treatment effect on different age groups, e.g., column 1 reports treatment effect on children that were below 2 months of age at baseline (thus, below 14 months at endline), while column 2 reports treatment effect on children that were between 2-4 months of age at baseline (thus, between 14-16 months at endline). Respective control group means at endline are provided right below each outcome (with standard deviations in brackets). All controls are listed under Table 3.

Table A3: Did mentally healthy mothers in the treatment arm responded biasedly about child outcomes?

	X: Trauma		Y: Depression	
	Without covariates	With covariates	Without covariates	With covariates
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
***A. Child's mental health*‡**				
Trauma severity	-0.059	-0.045	-0.104*	-0.087
	(0.064)	(0.061)	(0.060)	(0.058)
Depression severity	-0.104	-0.103	-0.113*	-0.110*
	(0.069)	(0.066)	(0.064)	(0.062)
Composite mental health index	-0.092	-0.083	-0.126*	-0.113*
	(0.069)	(0.066)	(0.065)	(0.062)
B. Child's development				
Communication skills	0.258***	0.249***	0.234***	0.217***
	(0.064)	(0.062)	(0.064)	(0.062)
Gross-motor skills	0.220***	0.211***	0.186***	0.176***
	(0.065)	(0.061)	(0.063)	(0.060)
Fine-motor skills	-0.007	-0.014	-0.029	-0.044
	(0.076)	(0.072)	(0.072)	(0.068)
Problem-solving skills	0.198***	0.195***	0.174***	0.165***
	(0.062)	(0.059)	(0.061)	(0.057)
Social skills	0.166**	0.175**	0.111	0.112
	(0.070)	(0.070)	(0.068)	(0.068)
Composite child development index	0.219***	0.215***	0.177**	0.164**
	(0.077)	(0.073)	(0.074)	(0.071)
Observations	2,228	2,223	2,589	2,584

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Vertical panel X (trauma) includes all participants from the control arm but, among the treated, only participants that experienced either deterioration or no improvement in trauma at endline. Similarly, vertical panel Y (depression) includes all participants from the control arm but, among the treated, only participants that experienced either deterioration or no improvement in depression severity at endline. Columns (1) and (3): treatment effects estimated without any baseline covariates. Columns (2) and (4): treatment effect estimated with all baseline covariates (as in equation 3). Covariates are listed under Table 3. For outcomes with ‡, negative coefficients imply more favorable outcomes.

Table A4: Robustness check for experimenter demand effect: using mothers' opinions on children's growth

	Treat	**Treat×Weight**	**Treat**	**Treat×Height**	**Treat**	**Treat×Growth**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
A. Children's mental health‡						
Trauma severity	-0.094*	0.200	-0.097*	0.402	-0.090	0.347**
	(0.055)	(0.156)	(0.055)	(0.296)	(0.055)	(0.146)
Depression severity	-0.120**	0.121	-0.123**	0.205	-0.120**	0.356**
	(0.059)	(0.152)	(0.059)	(0.272)	(0.059)	(0.149)
Composite mental health index	-0.121**	0.190	-0.125**	0.374	-0.118**	0.417***
	(0.058)	(0.155)	(0.059)	(0.294)	(0.058)	(0.148)
B. Children's development						
Communication skills	0.228***	-0.203	0.230***	0.069	0.222***	-0.175
	(0.059)	(0.133)	(0.059)	(0.186)	(0.059)	(0.153)
Gross-motor skills	0.176***	-0.252*	0.180***	-0.454**	0.170***	-0.191
	(0.058)	(0.139)	(0.058)	(0.212)	(0.057)	(0.157)
Fine-motor skills	-0.023	0.078	-0.021	0.146	-0.027	0.069
	(0.066)	(0.138)	(0.066)	(0.199)	(0.066)	(0.168)
Problem-solving skills	0.176***	-0.005	0.177***	0.018	0.172***	0.006
	(0.055)	(0.133)	(0.055)	(0.196)	(0.055)	(0.149)
Social skills	0.128*	-0.196	0.128*	-0.134	0.121*	-0.218
	(0.067)	(0.123)	(0.067)	(0.209)	(0.067)	(0.151)
Composite child development index	0.180***	-0.155	0.183***	-0.093	0.173**	-0.128
	(0.069)	(0.138)	(0.069)	(0.207)	(0.068)	(0.173)
Observations	2,840	2,840	2,840	2,840	2,840	2,840

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. Dependent variables are standardized indices, such that the control group has mean 0 and standard deviation 1. Columns 1, 3, and 5 report the coefficients on the treatment dummy with all covariates while also controlling for opinions about child's weight, height, and growth respectively. Opinions about child's weight, height, and growth are dummies, where it equals 1 if mothers are very satisfied with their children's weight, height, and growth respectively. Columns 2, 4, and 6 report the coefficient on the interaction between treatment and opinions about child's weight, height, and growth respectively. A statistically significant interaction coefficient suggests that treatment effect varies by mothers' options.

Table A5: Correlation of mental health between mothers and children

	Trauma of Children			Depression of Children		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
VARIABLES	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Panel A: At baseline					
Trauma of Mothers	0.188***	0.172***	0.201***			
	(0.027)	(0.033)	(0.033)			
Depression of Mothers				0.190***	0.186***	0.200***
				(0.048)	(0.058)	(0.072)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
Camp FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,493	1,705	1,788	3,493	1,705	1,788
R-squared	0.094	0.104	0.094	0.048	0.050	0.057
	Panel B: At endline					
Trauma of Mothers	0.246***	0.277***	0.215***			
	(0.028)	(0.038)	(0.039)			
Depression of Mothers				0.157***	0.173***	0.140***
				(0.031)	(0.041)	(0.044)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
Camp FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,798	1,382	1,416	2,798	1,382	1,416
R-squared	0.083	0.110	0.081	0.034	0.038	0.043

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. Dependent variables are standardized trauma (columns 1-3) and depression (columns 4-6) indices (same as in A1 and B1 panels in Table 3). Columns 1 and 4 report estimates of the entire sample, whereas the remaining columns report estimates by child's gender. Controls are listed under Table 3.

Table A6: Heterogeneous treatment effects, by camp abuse and duration in camp

		Victim of at least one camp abuse			Living in the camp for 2 years or less		
	Pooled	**Yes**	**No**	**Difference (β_3)**	**Yes**	**No**	**Difference (β_3)**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
A1. Mothers' mental health							
Trauma	-0.233***	-0.344***	-0.214***	-0.071	-0.315***	-0.143**	-0.201**
	(0.051)	(0.120)	(0.052)	(0.101)	(0.071)	(0.062)	(0.090)
Depression	-0.144***	-0.206**	-0.133**	-0.015	-0.207***	-0.073	-0.135
	(0.054)	(0.104)	(0.053)	(0.084)	(0.072)	(0.066)	(0.083)
A2. Mothers' wellbeing							
Happiness	0.117**	0.090	0.118**	0.016	0.154*	0.081	0.071
	(0.056)	(0.105)	(0.058)	(0.101)	(0.080)	(0.070)	(0.098)
Aspirations	-0.066	-0.160*	-0.047	-0.072	-0.062	-0.085	-0.049
	(0.062)	(0.096)	(0.068)	(0.105)	(0.090)	(0.075)	(0.107)
Belongingness	0.179***	0.322***	0.153***	0.175	0.126	0.205***	-0.098
	(0.057)	(0.110)	(0.058)	(0.111)	(0.077)	(0.076)	(0.102)
B1. Children's mental health							
Trauma	-0.096*	-0.057	-0.107*	0.130	-0.095	-0.110	0.022
	(0.055)	(0.110)	(0.058)	(0.101)	(0.076)	(0.071)	(0.096)
Depression	-0.122**	-0.210*	-0.103*	-0.074	-0.212**	-0.032	-0.158
	(0.059)	(0.119)	(0.061)	(0.111)	(0.092)	(0.059)	(0.107)
B2. Children's development							
Communication skills	0.229***	0.253**	0.230***	-0.025	0.209***	0.252***	-0.056
	(0.059)	(0.100)	(0.062)	(0.090)	(0.080)	(0.077)	(0.099)
Gross-motor skills	0.179***	0.156	0.186***	-0.087	0.137*	0.239***	-0.041
	(0.058)	(0.101)	(0.061)	(0.101)	(0.077)	(0.075)	
Fine-motor skills	-0.021	0.035	-0.023	-0.055	-0.039	0.012	-0.037
	(0.066)	(0.118)	(0.069)	(0.107)	(0.095)	(0.077)	(0.109)
Problem-solving skills	0.177***	0.037	0.206***	-0.249**	0.115	0.244***	-0.120
	(0.055)	(0.115)	(0.056)	(0.110)	(0.078)	(0.067)	(0.095)
Social skills	0.128*	0.209*	0.120*	-0.022	0.167*	0.099	0.102
	(0.067)	(0.121)	(0.071)	(0.120)	(0.093)	(0.077)	(0.106)
Observations	2,840	449	2,391	2,840	1,445	1,395	2,840

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. Columns 2-3 and 5-6 report estimates from split samples. Columns 4 and 7 report the coefficients on the interaction term from equation 5.

Table A7: Attrition and baseline characteristics

VARIABLES	**Only Baseline** (*Std. Dev.*)	N_{OB}	**Baseline & Endline** (*Std. Dev.*)	N_{BE}	**T-test** ***p*-values**	**RI** ***p*-values**
A: Mother & household characteristics						
Age	25.57	653	25.49	2,842	0.75	0.73
	(5.89)		(5.73)			
Whether receives voucher	0.49	653	0.51	2,842	0.48	0.48
	(0.50)		(0.51)			
Household size	5.22	654	5.26	2,845	0.70	0.72
	(2.00)		(1.98)			
Employed	0.03	653	0.03	2,842	0.83	0.81
	(0.16)		(0.16)			
Monthly income	0.44	18	0.45	74	0.99	0.96
	(0.51)		(0.50)			
Husband alive	0.96	654	0.97	2,845	0.11	0.14
	(0.20)		(0.17)			
Number of children	2.98	654	2.91	2,845	0.41	0.41
	(1.99)		(1.94)			
Attended school	0.71	654	0.74	2,844	0.0.26	0.25
	(0.46)		(0.44)			
Months in camp	25.07	654	25.75	2,845	0.06*	0.05*
	(10.85)		(14.58)			
Mother is the HH head	0.26	654	0.20	2,845	0.04**	0.03**
	(0.44)		(0.40)			
Household victimization (conflict)	0.15	654	0.16	2,845	0.99	0.99
	(0.13)		(0.12)			
Mother's victimization (camp)	0.01	654	0.01	2,845	0.93	1.00
	(0.05)		(0.04)			
HH victim of at least one conflict abuse	0.87	654	0.87	2,845	0.58	0.70
	(0.33)		(0.34)			
Mother victim of at least one camp abuse	0.15	654	0.16	2,845	0.74	0.71
	(0.36)		(0.37)			
B: Child characteristics						
Age	14.54	654	14.38	2,845	0.67	0.72
	(6.48)		(6.45)			
Gender	0.53	654	0.51	2,845	0.31	0.21
	(0.50)		(0.50)			
Whether elder siblings attend HPL	0.02	654	0.03	2,845	0.23	0.20
	(0.15)		(0.18)			
Child's victimization (camp)	0.01	654	0.01	2,845	0.45	0.49
	(0.06)		(0.05)			
Child victim of at least one camp abuse	0.05	654	0.04	2,845	0.39	0.39
	(0.21)		(0.19)			
Weight (kg)	8.75	654	8.60	2,845	0.19	0.23
	(2.26)		(2.15)			
Height (cm)	75.07	654	74.17	2,845	0.09*	0.11
	(10.08)		(9.74)			

Note: Column 'Only Baseline' reports averages of mothers/children that only took part in the baseline and column N_{OB} reports the corresponding sample size. Column 'Baseline & Endline' reports averages of mothers/children that took part in both baseline and endline surveys, and column N_{BE} reports the corresponding sample size. See the note under Table 3.4 for all variable descriptions. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A8: Attrition, by treatment

	Treatment	Control	Interaction
VARIABLES	**(1)**	**(2)**	**(3)**
Treatment			-0.014
			(0.213)
Age of mothers	0.000	0.001	0.001
	(0.002)	(0.003)	(0.003)
Treatment×Age of mothers			-0.000
			(0.004)
Household Size	-0.004	-0.025	-0.025*
	(0.006)	(0.015)	(0.015)
Treatment×Household Size			0.021
			(0.016)
Mother attended school	0.011	-0.058*	-0.058*
	(0.023)	(0.034)	(0.033)
Treatment×Mother attended school			0.069*
			(0.040)
Household spending	-0.000	0.000**	0.000**
	(0.000)	(0.000)	(0.000)
Treatment×Household spending			-0.000**
			(0.000)
Duration in the camp	0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Treatment×Duration in the camp			0.001**
			(0.001)
Mother receives voucher	-0.012	-0.038	-0.038
	(0.024)	(0.036)	(0.036)
Treatment×Mother receives voucher			0.026
			(0.044)
Husband is alive	-0.019	-0.056	-0.056
	(0.052)	(0.078)	(0.078)
Treatment×Husband is alive			0.037
			(0.093)
Family member stranded	-0.034	0.013	0.013
	(0.025)	(0.046)	(0.046)
Treatment×Family member stranded			-0.047
			(0.053)
HH victimization (conflict)	-0.090	0.080	0.080
	(0.068)	(0.155)	(0.154)
Treatment×HH victimization			-0.170
			(0.169)
Mothers' victimization (camp)	-0.076	0.071	0.071
	(0.193)	(0.305)	(0.304)
Treatment×Mothers' victimization			-0.147
			(0.360)
Mother is the HH head	0.047*	0.068	0.068
	(0.028)	(0.045)	(0.045)
Treatment×Mother is the HH head			-0.021
			(0.053)
Number of children	0.005	0.016	0.016
	(0.007)	(0.015)	(0.015)
Treatment×Number of children			-0.011
			(0.017)
Age of children	-0.000	-0.006	-0.006
	(0.003)	(0.004)	(0.004)
Treatment×Age of children			0.006
			(0.005)
Gender of children	0.002	0.015	0.015
	(0.016)	(0.021)	(0.021)

Treatment×Gender of children			-0.013
			(0.026)
Weight of children (kg)	-0.001	0.007	0.007
	(0.007)	(0.011)	(0.011)
Treatment×Weight of children (kg)			-0.008
			(0.013)
Height of children (cm)	0.001	0.004**	0.004**
	(0.002)	(0.002)	(0.002)
Treatment×Height of children (cm)			-0.003
			(0.003)
Child's victimization (camp)	0.032	0.037	0.037
	(0.172)	(0.208)	(0.207)
Treatment×Child's victimization (camp)			-0.005
			(0.269)
Observations	1,907	1,586	3,493
R-squared	0.007	0.032	0.056
Attrition rate	0.12	0.27	-
Joint p-value on interactions	-	-	0.19

Robust standard errors clustered at the block level are in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All columns present estimates using a linear probability model, where the dependent variable is attrition, a dummy variable that equals 1 if a mother did not participate in the endline survey and 0 if she participated in both baseline and endline surveys. The sample in column 1 is mothers/children in the treatment group and the sample in column 2 is mothers/children in the control group. Column 3 pools all sample together. We do not interact the treatment dummy with 'HH victim of at least one conflict abuse', 'Mother victim of at least one camp abuse', and 'Child victim of at least one camp abuse' because these indicators were derived from the 3 victimization indices. All variables were measured at the baseline. Overall attrition rate is roughly 19% (654 out of 3,499 mothers did not participate in the endline).

Table A9: Treatment effects: Additional bounds analysis

	Unadjusted	Most Extreme Bounds		2^{nd}-Most Extreme Bounds		More Bounds		
	Treatment Effects	Lower	Upper	Lower	Upper	0.05 SD	0.10 SD	0.25 SD
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A1. Mothers' mental health								
Trauma	-0.233***	-1.778***	0.792***	-1.079***	0.213***	-0.227***	-0.221***	-0.200***
	(0.051)	(0.148)	(0.095)	(0.085)	(0.050)	(0.038)	(0.038)	(0.038)
Depression	-0.144***	-2.420***	0.914***	-1.033***	0.231***	-0.126***	-0.120***	-0.099**
	(0.054)	(0.217)	(0.101)	(0.090)	(0.049)	(0.040)	(0.040)	(0.040)
A2. Mothers' wellbeing								
Happiness	0.117**	-0.426***	0.933***	-0.103*	0.585***	0.095**	0.089**	0.068
	(0.056)	(0.061)	(0.078)	(0.058)	(0.060)	(0.043)	(0.043)	(0.043)
Aspirations	-0.066	-1.095***	1.091***	-0.875***	0.635***	-0.054	-0.061	-0.082*
	(0.062)	(0.087)	(0.098)	(0.081)	(0.068)	(0.049)	(0.049)	(0.049)
Belongingness	0.179***	-0.414***	0.984***	-0.002	0.500***	0.176***	0.169***	0.148***
	(0.057)	(0.063)	(0.081)	(0.057)	(0.058)	(0.046)	(0.046)	(0.046)
B1. Children's mental health								
Trauma	-0.096*	-2.089***	0.823***	-1.182***	0.316***	-0.094**	-0.087**	-0.066
	(0.055)	(0.189)	(0.090)	(0.107)	(0.053)	(0.042)	(0.042)	(0.042)
Depression	-0.122**	-2.574***	0.893***	-1.244***	0.485***	-0.111**	-0.105**	-0.084*
	(0.059)	(0.235)	(0.101)	(0.113)	(0.068)	(0.045)	(0.045)	(0.046)
B2. Children's development								
Communication skills	0.229***	-0.454***	0.867***	-0.420***	0.762***	0.226***	0.220***	0.201***
	(0.059)	(0.069)	(0.070)	(0.068)	(0.063)	(0.048)	(0.048)	(0.048)
Gross-motor skills	0.179***	-0.408***	0.850***	-0.049	0.659***	0.174***	0.168***	0.149***
	(0.058)	(0.062)	(0.072)	(0.058)	(0.068)	(0.046)	(0.046)	(0.046)
Fine-motor skills	-0.021	-0.608***	0.581***	-0.205***	0.405***	-0.015	-0.021	-0.040
	(0.066)	(0.068)	(0.072)	(0.064)	(0.072)	(0.053)	(0.053)	(0.053)
Problem-solving skills	0.177***	-0.400***	0.720***	-0.021	0.579***	0.178***	0.172***	0.153***
	(0.055)	(0.061)	(0.062)	(0.055)	(0.062)	(0.044)	(0.044)	(0.044)
Social skills	0.128*	-0.542***	0.829***	-0.087	0.576***	0.113**	0.107**	0.088*
	(0.067)	(0.072)	(0.075)	(0.068)	(0.069)	(0.052)	(0.052)	(0.052)

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Column (1) reports unadjusted treatment effects, same as in Table 3. Columns 2-5 report treatment effects with extreme bounds following Horowitz & Manski (2000). Columns 6-9 report treatment effects with moderate bounds following Kling et al. (2007) and Karlan & Valdivia (2011). All specifications control for baseline characteristics and standard errors are clustered at the unit of randomization.

Table A10: Spillover effects on nearby blocks

	Mother's trauma severity					Mother's depression severity				
	Adj	**Adj No.**	**Adj-%**	**200m**	**400m**	**Adj**	**Adj No.**	**Adj-%**	**200m**	**400m**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Treat	-0.346***	-0.365***	-0.312***	-0.303***	-0.351***	-0.103	-0.157	-0.074	-0.147	-0.158
	(0.127)	(0.109)	(0.117)	(0.100)	(0.110)	(0.123)	(0.107)	(0.114)	(0.102)	(0.105)
Adjacent	-0.129					-0.066				
	(0.143)					(0.135)				
Treat×adjacent	0.159					-0.042				
	(0.149)					(0.149)				
No. of adjacent		-0.073					-0.027			
		(0.071)					(0.074)			
Treat×No. of adjacent		0.135*					0.026			
		(0.080)					(0.092)			
% of treat adjacent			-0.119					0.153		
			(0.335)					(0.333)		
Treat×% of treat adjacent			0.332					-0.246		
			(0.375)					(0.400)		
Treated in 200m radius				-0.039					-0.025	
				(0.065)					(0.070)	
Treat×Treated in 200m radius				0.079					0.019	
				(0.075)					(0.093)	
Treated in 400m radius					-0.047					-0.007
					(0.069)					(0.073)
Treat×Treated in 400m radius					0.118					0.032
					(0.079)					(0.089)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Camp FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,801	1,801	1,788	1,801	1,801	1,801	1,801	1,788	1,801	1,801
R-squared	0.050	0.052	0.049	0.049	0.051	0.038	0.036	0.037	0.036	0.036

Robust standard errors clustered at the block level are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. The outcome variable in columns 1-5 is trauma and that in columns 6-10 is depression severity. Both outcomes are standardized indices, such that the control group has mean 0 and standard deviation 1. 'Treat' is a dummy that equals 1 if the block is treated and 0 if control; 'Adjacent' is a dummy that equals 1 if a block has at least 1 adjacent block that is treatment and 0 otherwise; 'No. of adjacent' is the number of adjacent treatment blocks; '% of treat adjacent' is the number of adjacent divided by the total number of adjacent blocks; 'Treated in 200m radius' and 'Treated in 400m radius' are the number of treatment blocks within the 200 and 400 meter radius of each block. This information is only available on roughly 1,800 individuals, which explains the smaller sample sizes.

B Appendix: Details on over-the-phone sessions

To accommodate over-the-phone sessions, session duration and structure were carefully revised by experts at the BRAC Institute of Education and Development. We explain these changes in detail below:

- **Over-the-phone sessions:** Over-the-phone sessions replaced the usual face-to-face sessions after 24 weekly sessions. Initially, BRAC only had mobile phone numbers of 42% of enroled mothers or their spouses. To collect contact numbers from the remaining households, mother volunteers, project assistants, and community leaders went door-to-door and gathered contact numbers from an additional 45% of enroled mothers.[6] If a mother did not have any mobile phone, they collected the contact number of another household member (which were typically their spouses). In total, 87% of enroled mothers had mobile phones and mobile phone ownership does not differ by treatment groups (T-test: $p > 0.10$).[7] Contact numbers from 13% of enroled mothers (similar across treatment groups) could not be collected because neither them nor their household members owned any mobile phones. Therefore, 13% of the mother-child pairs in the treatment group could not be treated over-the-phone due to the pandemic. During the preparation period, MVs were also remotely trained for over-the-phone treatment delivery by the same experts that trained them for face-to-face treatment delivery. Eventually, over-the-phone sessions started in mid-May 2020 and ended in late-September 2020, following the completion of 20 weekly sessions.
- **Duration:** The duration of the over-the-phone sessions had to be adjusted to make it feasible for telephone conversations. Thus, hour-long sessions had to be curtailed to 20-minutes-long sessions. Analogous to face-to-face sessions, telephone sessions were also conducted on a weekly basis. However, over-the-phone sessions were conducted on a one-to-one basis.
- **Content:** Face-to-face sessions were more interactive, as it had a combination of discussion and sharing with MVs and other participants, and play activities for both mothers and children. However, play activities, breathing exercises in groups, interaction with children, and various group discussions could not be conducted during over-the-phone sessions. Therefore, over-the-phone sessions excluded these activities and instead focused on: (1) Step 1—feedback on homework for 3 minutes; (2) Step 2—discussion on well-being and sharing with MVs for 8 minutes; (3) Step 3—childcare discussions and practices for 8 minutes; and (4) Step 4—homework for 1 minute. MVs also encouraged mothers to continue doing the physical activities and play activities with children that they had learned from face-to-face sessions. Finally, during Steps 2 and 3, information on COVID-19 preventive measures, such as social distancing, hand-washing, and coughing/sneezing etiquette, was also added to the telecounseling modules to help the mothers and their families keep safe and informed during the pandemic.

[6] Field staff strictly followed COVID-19 health guidelines while collecting mobile phone numbers from households.

[7] Also add balance check table for appendix: phone owner vs no phones.