I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16961

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

Asad Islam
Tanvir Ahmed Mozumder
Tabassum Rahman
Tanvir Shatil
Abu Siddique

APRIL 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16961

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

Asad Islam
Monash University

Tanvir Ahmed Mozumder
BRAC Institute of Governance & Development

Tabassum Rahman
University of Melbourne

Tanvir Shatil
BRAC Institute of Governance & Development

Abu Siddique
Royal Holloway University of London and IZA

APRIL 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

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

Forced displacement is a major driver of mental health disorders among refugees globally. The mental well-being of adult refugees, particularly mothers, is widely recognized as a crucial determinant of their children's psychological health and development. In this study, we conducted a randomized controlled trial (RCT) to examine the effectiveness of a multifaceted psychosocial program in improving the mental health of refugee mothers, and fostering growth and development among children under the age of two. Collaborating with BRAC, we conducted a cluster RCT involving 3,500 Rohingya mother-child pairs in refugee camps in Bangladesh. Participants received weekly psychosocial support for 44 weeks, facilitated by trained peer volunteers. The program included psychoeducation and parenting guidance for mothers, as well as interactive play activities for both mothers and children. The intervention proved largely successful, resulting in: (i) reductions in the psychological trauma and depression severity among both mothers and children, (ii) improvements in children's communication, gross-motor, and problem-solving skills, and (iii) reductions in the prevalence of stunting and severe stunting among children. At a cost of approximately $1 per dyad per session, the intervention has demonstrated cost-effectiveness and is currently being scaled-up in Bangladesh's refugee camps, benefiting around forty thousand mother-child dyads.

JEL Classification: I15, J15, O12, O15

Keywords: mental health, forced displacement, early childhood development, refugees

Corresponding author:
Abu Siddique
Department of Economics, Royal Holloway University of London
Great Britain

E-mail: abu.siddique@rhul.ac.uk

* We thank Abi Adams-Prassl, Sule Alan, Daniel Bennett, Sonia Bhalotra, Janet Currie, Narayan C Das, Fabian Dehos, Adeline Delavande, Amrita Dhillon, Steven Dieterle, Eliana La Ferrara, Soledad Giardili, Sebastian Goerg, John Hoddinott, Joo Young Jeon, Anett John, Patricia Justino, Hanna Kienzler, Tatiana Kornienko, Kostas Matakos, Imran Matin, Imran Rasul, Mushfiq Mobarak, Karthik Muralidharan, Ana Nuevo-Chiquero, Vikram Patel, Catherine Porter, Yasuyuki Sawada, Frank Schilbach, Smriti Sharma, Saurabh Singhal, and Andrea Smurra for many helpful comments and suggestions. We also thank seminar participants at Australian National University, Adelaide, Bath, Derby, Duisburg-Essen, Edinburgh, Foundations UK, IFS-UCL-LSE/STICERD Development, Indian Statistical Institute-Delhi, King's College London, Lancaster, Monash, National Graduate Institute for Policy Studies (GRIPS), Newcastle, NOVAFRICA, Osaka, Royal Holloway, Technical University of Munich, Tokyo, Washington, Western Australia, and conference participants at the 46th BREAD-MIT conference, NBER Summer Institute 2022 (development session), ASSA 2022, CEPR/LEAP Workshop in Development Economics, and other conferences for many helpful comments. We also thank Erum Marium and Sakila Yesmin from the BIED and BRAC fieldworkers in Rohingya refugee camps for extensive support and collaboration for the entire project. The project received funding for the intervention from the LEGO Foundation and the Porticus Foundation, with research and evaluation funding provided by the Open Society Foundation. The study has received research ethics approval from BRAC University (2019-028-ER), and is registered at the AEA RCT Registry (AEARCTR-0004516)

1 Introduction

> *"When I try to sleep, I imagine what the military has done to me. I feel like they are coming, chasing, and shooting me... When I am in bed, the imagination of the torture appears in my mind."*
>
> Rashida Begum, a Rohingya woman (Fortify Rights, 2020).

The global refugee crisis, caused by conflict, persecution, natural disasters, and famine, has resulted in the displacement of over 100 million people worldwide (UNHCR, 2022). Refugees are highly susceptible to developing common mental disorders, such as depression, trauma, and anxiety, due to their exposure to stressful life events like violence, separation, financial strain, and uncertainty (Steel et al., 2009; Song & Teichholtz, 2019). While social science and public health research have long sought to understand how stressors affect the mental health of refugees (Stillman et al., 2009; Miller & Rasmussen, 2010), relatively little research has been done to understand low-cost mitigation policies suitable for poor humanitarian contexts.

In this paper, we address this gap by presenting evidence from a large cluster randomized controlled trial to evaluate the impact of a multifaceted psychosocial support program that was designed to improve the mental health of refugee women in Bangladesh (mothers, henceforth) and socioemotional, physical, cognitive, and anthropometric development of their children under the age of two. At the time of implementation, it was one of the world's largest interventions on the mental health of refugees. Our intervention targeted mothers and children, a highly vulnerable group among forcibly displaced populations (UNHCR, 2021a), due to the heightened risks refugee women face of complex trauma from violence and abandonment (Shishir, 2022), and the increased susceptibility of refugee children to malnutrition and diarrhoeal diseases from inadequate care (Hossain et al., 2019).

The economic cost of mental health problems is very high (Bloom et al., 2012), particularly in low- or lower-middle-income countries (Collins et al., 2011). One way mental health interacts with economic outcomes is through decision-making and choices. Mental health affects cognitive functions, such as memory, attention, logical reasoning, and problem-solving, which are essential for making informed choices (Stewart & Ware, 1992). Poor mental health, like depression or anxiety, disrupts these functions, leading to suboptimal decisions. Poor mental health can also skew an individual's time preferences, making them likely to prioritize immediate rewards over future benefits (Bayer & Osher, 2018). Therefore, in our context, poor mental health can limit the capacity of mothers to provide adequate care and support to their children, as well as affect their willingness to engage in activities with children that promote their cognitive, social, and emotional development, such as reading, playing games, or chanting rhymes. This, as a result, may affect the quality and quantity of parental input that are essential during early childhood (Carneiro et al., 2023). Moreover, children who grow up in households with mentally unhealthy mothers may also experience poor mental health themselves (Bütikofer et al., 2024). The economic costs of poor mental health and malnutrition during childhood are also substantial in the long run, as they can impede human capital accumulation, cause poor mental health in adulthood, and perpetuate the cycle of poverty (Heckman et al., 2006; Currie, 2009; Adhvaryu et al., 2019; Ridley et al., 2020; Walker et al., 2022). Therefore, the impact of poor mental health of mothers on children's development in poor and humanitarian contexts can be far-reaching and multidimensional, highlighting the importance of addressing maternal mental health in these contexts.

As a potential solution to the mental health issue among refugees, the BRAC Institute of Education and Development (BIED) pioneered a program called the home-based Humanitarian Play Lab (HPL). This program was implemented on a sample of 3,500 Rohingya refugee mother-child dyads located in refugee camps in Bangladesh. The target was the Rohingya people—a severely persecuted ethnic and religious minority from Myanmar. The 2017 incidence of mass genocide and community violence in Myanmar caused a mass displacement of about 750,000 Rohingya people to Bangladesh, where currently a million Rohingyas live as refugees in confined camps. In the camps, the mental health of Rohingya women and children is alarmingly poor, and acute malnutrition, anemia, and stunting are very high among children (Hossain et al., 2019). Moreover, the legal status of Rohingya refugees in Bangladesh does not allow their social inclusion, participation in employment, or mobility out of camps, which imposes an additional mental toll on the refugees.

Against this background, the home-based HPL intervention was designed to provide both mental support for mothers and to facilitate quality interactions with their children. The intervention included psychoeducation and support for parenting, including counseling on psychosocial stimulation, provided to mothers by trained community peers who served as volunteers. Psychoeducation is an established psychosocial support tool that integrates light-touch psychotherapeutic and educational interventions to help people cope with common mental health problems (American Psychological Association, 1995; Fusar-Poli et al., 2021).[1] Together with psychoeducation and parenting support, mothers and children also engaged in culturally appropriate play activities during treatment sessions. Participants in the control arm also attended *unstructured* social gatherings on a weekly basis but they were not given the psychoeducation treatment. The treatment was provided weekly for a year, from October 2019 until September 2020, through 44 weekly sessions. Due to the COVID-19 pandemic, the last 20 sessions were delivered over mobile phones.

Thirteen months after the intervention began or one month after the intervention ended, we found that mothers who received psychosocial support experienced a 0.23 standard deviation (SD) reduction in symptoms of psychological trauma and a 0.14 SD reduction in depressive symptoms, indicating a direct impact on common mental health challenges faced by refugees. Among the mothers who were identified to have psychological trauma and depression at baseline, we observed a sizeable improvement in their mental well-being following the intervention. In addition, treated 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. These results are also important because happiness is a key factor of well-being, reflecting positive emotions and outlook on life. Additionally, a sense of belonging promotes integration into the host community, which is essential for refugees who often experience isolation and social exclusion after being displaced. We, however, did not find any noticeable impact on their aspirations for the future. This could be explained by our program's focus on addressing present needs for mental health support. Future aspirations might require a different approach focusing on goal setting, financial planning, or educational opportunities, which participants did not receive. Furthermore, mentally unhealthy mothers who received the treatment caught up to, and often surpassed, the mental health of the

[1] Psychoeducation educates people in mental hardship about the possible reasons for their distress and simple ways of addressing it. It also facilitates discussion and sharing of various positive and negative feelings with others, which helps people identify the challenges they are facing and their personal coping abilities (Cuijpers et al., 2009; Donker et al., 2009). Given its simplicity, psychoeducation can be easily delivered by non-experts from poor settings with limited educational backgrounds.

'mentally healthy' mothers in the control group following the intervention. Thus, the intervention was largely successful in lifting refugee mothers out of psychological distress.

Children in the treatment arm also experienced reductions in psychological trauma (0.10 SD) and depressive symptoms (0.12 SD) relative to children in the control arm, but these differences are only significant at 5%–10% levels. This finding aligns with research demonstrating the effectiveness of psychosocial stimulation (play activities between mothers and children) in promoting positive growth outcomes in children (Grantham-McGregor et al., 1991), and short- and long-term mental health benefits, such as reductions in anxiety and depression (Nahar et al., 2009; Walker et al., 2006). On the other hand, research also suggests a positive correlation between children's mental health and their parents' mental health (Bütikofer et al., 2024). Therefore, the improvements in mothers' psychological well-being observed in our intervention might have also positively impacted the children's mental health. Due to the bundled nature of our intervention, we cannot tease out these two potential channels.

We also find that the intervention improved treated children's communication skills (speech and language development) by 0.23 SD, gross-motor skills (physical activities and whole-body movements) by 0.18 SD, and problem-solving skills (learning to play with toys and solve puzzles) by 0.18 SD. We also observe a marginal improvement in children's personal-social skills (caring for themselves and interacting with others) by 0.13 SD, $p < 0.10$). These results can be explained by the intervention's content. Sessions included free play with age-appropriate toys, encouraging children's curiosity, communication, and muscle movements. Activities like peekaboo and toy hunts directly target communication, gross motor, and problem-solving skills. The program also emphasized psychosocial stimulation and play using household items. These activities likely contributed to the gains in problem-solving and gross motor skills by encouraging creativity and physical movement. However, we failed to detect any statistically sizable impacts on children's fine-motor skills development (small muscle movements). The intervention's emphasis on playing and moving larger muscles likely explains the significant gains in children's gross motor skills but not fine motor skills. Because these results are based on survey responses, we address concerns about social desirability bias and experimenter demand effects in multiple ways, including using the Marlowe-Crowne scale (Dhar et al., 2022).

We also find that children in the treatment arm experienced a large increase in height-for-age z-score by 0.52 SD (19% or 1.58 centimeters taller), which also translates to a 7 percentage points (or 10%) reduction in stunting and a 13 percentage points (or 22%) reduction in severe stunting (skeletal growth retardation). This aligns with the findings of Nahar et al. (2009), who showed that even two weeks of psychosocial stimulation can lead to improvements in growth outcomes (0.39 SD units) among children in Bangladesh. However, other studies have reported mixed results (Super et al., 1990; Hamadani et al., 2006). These mixed findings suggest that interventions leading to anthropometric improvements in children might be primarily driven by changes in mothers' behavior, such as improved childcare practices (Nahar et al., 2009). It is important to note that due to the COVID-19 pandemic lockdown, anthropometric enumerators could not measure children's height in person. Instead, mothers were asked to measure their children's height over mobile phones using their right hand and index finger, with the 'hand-finger' units being later converted to centimeters following Asadujjaman et al. (2019). These measurements were partially validated by community

leaders (known as *majhee*) who randomly visited participating mothers.[2] While this partial validation provides some support for the reported results, it is crucial to interpret the anthropometric findings with a degree of caution.

To understand if other potential mechanisms explain our findings, we estimate the impact of the intervention on several intermediate outcomes. Along with the psychotherapeutic aspect, the educational aspect of psychoeducation encompassed advice and suggestions on self-care (such as healthy eating, exercise, and adequate sleep), family communication, and social connections for mothers. We, however, did not find any statistically significant evidence to support these potential mediators for mothers' outcomes, implying that mothers' mental health might have benefited directly from the psychotherapeutic aspects (i.e., psychoeducation) of sessions.

On the other hand, some elements of the parenting advice seem to be potential mechanisms for children's outcomes. We found strong evidence that the intervention increased mothers' self-reported daily interaction with their children by approximately 1.5 hours, but had no effects on fathers' involvement, suggesting that maternal time investment in children could be a potential channel. Additionally, treated mothers were less likely to allow their children to play or walk barefoot (reduces the risk of hookworm infections and exposure to various bacteria and fungi), and also less likely to engage in negative parenting, which could be other potential channels for their children's development. These changes in behavior suggest that the intervention positively shifted how mothers interacted with and cared for their children. The intervention's focus on mothers' mental well-being may have also made them more receptive to the parenting advice. Improved mental health can increase a mother's ability to understand the rationale behind better childcare practices and potentially reduce the effort cost associated with implementing the advice. These explanations align with the findings of Carneiro et al. (2023), which shows that parenting programs can improve parenting behaviors and the home environment—both crucial factors for early childhood development. Similarly, research suggests a strong relationship between poor maternal mental health and child malnutrition, as poor mental health can affect the "emotional quality of childcare" and that improving maternal mental health can lead to reductions in stunting and wasting among children (Rahman et al., 2004).

Finally, we examine heterogeneity in treatment effects using machine learning following Chernozhukov et al. (2020). Regarding the mental well-being of mothers, those with poor mental health at baseline, high exposure to violent conflict in Myanmar, and more abuse in refugee camps had the most substantial benefits from the intervention. There is also weaker evidence that older and illiterate mothers had the greatest benefits. In terms of children's skills development and anthropometric outcomes, older children saw the most improvement across all dimensions. However, there was no difference in stunting between boys and girls. We also do not observe heterogeneity in child development and anthropometric outcomes based on baseline mental health.

Overall, the intervention was largely successful and cost-efficient, with a cost of approximately USD 45 per mother-child dyad for 44 weekly sessions. BRAC Bangladesh is currently expanding the program in the Rohingya refugee camps. To date, over forty thousand mother-child pairs, including some from the control group, have already benefited from the program.

[2]We discuss this in details in section 3.3.1. Weight was also measured as the 'best guessed' weights of children by mothers. However, weights could not be validated later by *majhees* and are entirely subjective and rather noisy, which is why we dropped weight-for-age z-score and weight-for-height z-score outcomes (both pre-registered) from this paper. These results can be made available upon request.

Our paper contributes to multiple strands of literature. First, it adds to our understanding of the relationship between psychotherapy, mental health, and economic outcomes. Psychotherapy methods like cognitive behavioral therapy (CBT), behavioral activation, problem-solving therapy, and psychoeducation are effective in treating mental health issues across various contexts (Rahman et al., 2008; Cuijpers et al., 2009), leading to improved decision-making and long-term economic benefits of targeted people (Bhalotra et al., 2020; Barker et al., 2022; Bhat et al., 2022), even in post-conflict settings (Rahman et al., 2019). Our intervention incorporates psychoeducation, a proven light-touch method focusing on psychotherapy and mental health awareness (Cuijpers et al., 2009).[3] Cuijpers et al. (2009), a meta-analysis on psychoeducation, shows that psychoeducation therapy can reduce the risk of getting major depression by 38% and can improve depressive symptoms by 0.28 SD. They do not find evidence of psychoeducation being less effective than other psychotherapy treatments.

Additionally, our research builds on the body of work surrounding early childhood development (ECD) programs focused on psychosocial stimulation through play, parenting counseling, or a combination of both (Grantham-McGregor et al., 1991; Singla et al., 2015). There is strong evidence that quality ECD interventions have lasting positive impacts on human capital and later-life economic outcomes (Heckman et al., 2013; Gertler et al., 2021; Attanasio et al., 2022). More broadly, our study also relates to the literature on the importance of early-life interventions on child development and human capital accumulation (Miguel & Kremer, 2004; Alan et al., 2021; Carneiro et al., 2021). We extend this line of literature by providing further insights into the importance of early-life interventions to mitigate the adverse effects of childhood trauma caused by conflict or displacement.

Our study offers several unique contributions. We integrate the various strands of literature discussed above into a large-scale, well-powered field experiment. Limited evidence exists on similar interventions' impact on both maternal mental health and child development. Two exceptions are Singla et al. (2015) and Baumgartner et al. (2021), who used cluster-RCTs in Uganda and Ghana, respectively, with smaller samples but mixed results. Our study addresses this gap by evaluating a similar intervention with a more marginalized, at-risk population, furthering our understanding of when these interventions are effective. Our approach also avoids costly baseline screening processes to target both those with current and potential future mental distress. This approach is designed for scalability and to potentially reduce mental health stigma. Therefore, this study was designed to be scalable by removing the screening process and training community volunteers (i.e., refugee women) as program facilitators. Additionally, our study uses a placebo control group to estimate the causal impact of the program, which is a unique aspect compared to most mental health trials. Furthermore, well-powered psychosocial programs for refugees are uncommon, and our study now addresses this gap in the literature. Lastly, our study offers novel data on the connection between psychoeducation and early childhood development outcomes. To better understand how our study and sample differ from existing literature, we have also summarized the most relevant studies on mental health interventions in Table C1, Appendix C.

[3] Informational and light-touch psychoeducation therapy (delivered remotely) have also been proven to be effective in reducing depressive symptoms, stress, and anxiety among people in isolation (Vlassopoulos et al., 2024).

2 The context

> *"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).

> *"Still traumatised after fleeing violence in Myanmar, Nazima Begum is struggling to breast-feed her seven-month-old son. Her story is all too common among the hundreds of thousands of women who have taken refuge in Bangladesh."*
>
> Ford (2018) on Nazima's struggle with mental health and breastfeeding.

> *"At night time, I have to keep him beside me all the time. Sometimes he gets convulsions. He makes a big sound when he gets a convulsion."*
>
> Rohima on her son's struggle with trauma (Save the Children, 2019).

The Rohingya people of Myanmar (previously Burma) are an ethnic, linguistic, and religious minority in Myanmar. 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 and in late 2012. A new wave of violence against the Rohingya people spurred in 2017, also known as 'ethnic cleansing' by the Burmese military, forced the majority of Rohingyas to seek refuge in neighboring Bangladesh (Beyrer & Kamarulzaman, 2017). During this incident, about 24 thousand Rohingyas were killed, 18 thousand women and girls were raped, 34 thousand were thrown in the fire, 114 thousand were severely beaten, and over 100 thousand households were burned down or vandalized (Habib 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.

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. Moreover, these camps consist of 31% vulnerable families, such as separated children and families with single mothers, with at least one protection vulnerability. According to WHO Situation Report (2017), around half of the Rohingya children in refugee camps in Bangladesh are malnourished, underweight, and suffering from anemia, and 25% of the children under 5 have acute malnutrition. Moreover, 38% of children have stunted growth—very close to the WHO critical health emergency threshold of 40% (Save the Children, 2018)—and over 80,000 children have severe mental distress, which is one in every five children in the camps (Save the Children, 2019). Over 30 thousand infants are born every year in camps that require quality nurturing from mothers and health experts (Tayeb, 2021a).

Immediately after fleeing Myanmar, over 80% of Rohingya women reported having depressive and emotional distress symptoms, and 60% had post-traumatic stress disorders (Fortify Rights, 2020). In the refugee camps, gender-based violence is very common, where most violence is initiated

by either intimate partners, relatives, or other camp members (Beech, 2017). Moreover, refugees cannot be employed, start new income-generating activities, or send their children to schools outside the camps due to legal restrictions. Therefore, they rely entirely on government support, foreign donors, and humanitarian agencies for food, healthcare, and shelter. Camps are also dense, with about 90 thousand people living in one square kilometer. Recent estimates also show that it will take 12 years if the Bangladeshi government repatriates 300 Rohingyas every day (Tayeb, 2021a).

3 The experiment

3.1 The home-based HPL program

The program. BRAC developed a psychosocial program called the "home-based Humanitarian Play Lab (HPL)" with the aim to improve the well-being of Rohingya mothers and their children under 2. With support from psychologists and early childhood experts from the BRAC Institute of Education and Development (BIED), and various external experts, this low-cost program was developed to run for 44 weeks through weekly sessions, to be delivered in a home setting by non-experts. This program was created as an urgent measure for persecuted and displaced Rohingya mothers and children, with the aim of scaling it after evaluating its impact.

The HPL program includes three components: (i) *Psychoeducation* (Cuijpers et al., 2009), aimed at helping Rohingya mothers cope with mental distress and trauma through education about their challenges and ways to address them, resulting in a better understanding of their coping abilities, strengths and weaknesses, and increased mental peace (American Psychological Association, 1995; Lukens & McFarlane, 2004); (ii) *Parenting support*, emphasizing the importance of childcare and early-childhood stimulation through play activities; and (iii) *Play activities* for mothers and children during sessions, including free-play with age-appropriate toys.

The HPL program was delivered by trained Rohingya refugee women (known as mother volunteers or MV) from the same neighborhood as the participants.[4] The program was provided weekly to small groups of participating mothers at the MV's home, with each 60-minute session led by the MVs who were trained by mental health and early-childhood experts, and received support from psychosocial experts when needed.

Session procedure. Each session was broken down into four steps: **(1) Greetings** (15 minutes) involved greetings and breathing exercises to relax participants, as well as a discussion of the previous week's homework. This step was identical in every session. **(2) My well-being** (20 minutes) covered psychoeducation and well-being advice for mothers, including self-care (such as healthy diet, the importance of sleep, nurturing hobbies, etc.), positive thinking, sharing positive and negative feelings with other participating mothers and MV, emotional development, play (e.g., hole tarp, bank-a-ball, etc.) and art activities for mothers, and free-play activities with age-appropriate toys for children. Topics varied per session. **(3) Baby's growing up** (20 minutes) offered parenting advice (e.g., spending quality time, timely feeding, nutrition, ways to massage a baby for better sleep, etc.) and psychosocial stimulation. Mothers were taught how they can play with their children with various household items, such as using a pillow, handkerchief, etc. Mothers also engaged

[4] MVs were hired by BRAC program managers and camp-in-charges based on their level of education, fluency in Bangla and Rohingya languages, and field management skills. Priority was given to women who knew how to read and write and were willing to set up sessions at their homes.

in play activities with their children during this step (e.g., peekaboo, toy hunt, counting fingers, etc.). Topics also varied per session. **(4) Homework** (5 minutes) assigned weekly tasks based on topics discussed.

The HPL program was randomly provided to selected mother-child pairs in the treatment group, while the control group participated in weekly *unstructured* social gatherings (thus, there was no curriculum, structured discussions, or MV to facilitate psychoeducation) on a weekly basis. This allows us to disentangle the effect of the program from the effect of attending social gatherings. All sessions were conducted in the local *Rohingya* language. The full curriculum (translated into English) is available here.

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. This intervention was not stopped after 24 in-person sessions for two reasons: first, experts from BIED recommended completing the entire curriculum of 44 sessions; and, second, due to humanitarian reasons, as the Covid-19 lockdown and uncertainty were likely to impose further mental toll on these vulnerable refugees.

The HPL program was adapted for over-the-phone sessions by experts from BIED, with revised duration and structure. Only individual sessions were conducted over the phone by the same MVs as group sessions (thus, play activities, group activities, and group discussions could not be conducted), each lasting 20 minutes. 87% of enrolled women had access to mobile phones, with the remaining 13% able to borrow phones from camp managers or block-*majhees* (leaders of each block).[5] The control group, however, did not receive placebo calls or engage in *unstructured* social gatherings. There were 20 weekly phone sessions. In section 3.6, we also show that the characteristics of women that initially had mobile phones are very similar to the characteristics of those that did not.

Timeline. Figure 1 shows the intervention timeline. The program began in October 2019 and ended in September 2020, with a temporary halt in March 2020 due to the Covid-19 lockdown. Over-the-phone support replaced face-to-face sessions in May 2020, lasting until September 2020. Baseline data were collected from July to September 2019, while endline data were collected over the phone throughout October 2020. No midline data was collected prior to phone sessions due to logistical constraints.

3.2 Sampling

Each refugee camp in Cox's Bazar consists of many blocks, 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. 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 (55%) and 114 were assigned to the control group (45%). Figure A1 in Appendix

[5]Note that participants were not forced to borrow mobile phones. Every week, prior to a scheduled session, *majhees* went to participants' doors and offered them their mobile phones for the session. After about an hour, *majhees* went back to collect the mobile phone at the door. Qualitative feedback from *majhees* suggests that mothers were not reluctant to borrow mobiles, but rather were very enthusiastic. Note also that social distancing measures were strictly followed and disinfectants provided by BRAC were applied on mobiles after each use.

A shows a camp map and blocks therein, highlighting the treatment and control blocks.

Within each block, we randomly created two groups, where each group attended an MV's home throughout the year for the sessions. We had a total of 226 groups in treatment and 191 in control blocks (one MV per group). For each session, we randomly invited roughly 7 mother-child dyads. From BRAC's list of Rohingya households, project assistants and MVs randomly visited households that met the selection criteria|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 HPL program. In case a mother had multiple children within this age category, we randomly selected one child for the intervention. A total of 3,499 mother-child dyads were enrolled to participate in this program. Only mothers in the treatment arm received our weekly treatment, while mothers in the control arm participated in *unstructured* (or unsupervised by an MV) social gatherings that did not involve psychosocial support or play activities.

3.3 Data

3.3.1 Data collection

The baseline data was collected in person by BRAC enumerators. The endline was conducted over mobile phones due to Covid-19 restrictions. Enumerators|both females and males|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. Our baseline questionnaires were divided into three broad parts: (i) socioeconomic background; (ii) mother's characteristics and adverse life experiences; and, (iii) adverse life experiences of children and age-specific questions on the skills development of children. At endline, only outcomes and potential mechanisms were collected. All survey questions were answered by mothers.

Trained anthropometric enumerators collected children's height measurements at baseline using infantometers. However, due to Covid-19 restrictions, at endline, they instructed mothers over the phone to measure their children's height using their right hand and index finger. Although this method using 'hand' and 'finger' units is outdated, it was the only way we could measure anthropometric outcomes during the pandemic without risking the health of participants, enumerators, and other Rohingyas in the camp.[6] Mothers reported the measurements in hand and finger units, which were later converted to centimeters following Asadujjaman et al. (2019). A validation process was conducted by block-*majhees* by randomly visiting participating mothers and asking them to demonstrate the measurement technique (about 20% of the total sample, or 2-3 mothers per block), marking it as correct or incorrect if it matched or differed from the initial measure.[7] But, among the randomly validated 20%, there were no large discrepancies, possibly because height is something factual and mothers were aware that its authenticity could be easily validated by BRAC. It could not be validated for all participants due to logistical constraints during the lockdown.

At baseline, a total of 3,499 mothers were surveyed: 1,911 in treatment and 1,588 in control. At endline, 2,845 mothers were surveyed (using mobile phones), 1,679 in treatment, and 1,166

[6]'Hand' length is the length between the mid-point of the wrist's distal transverse crease and the tip of the middle finger, and 'finger' width is the width of the index finger (see Figure A6 in Appendix A).

[7]Only 11 mothers (1.9% of revisited) made errors, e.g., using middle fingers for 'finger' lengths. Our results on height remain robust even when we drop 2% of the maximum gains in height in the treatment group, assuming the maximum 2% was due to measurement errors.

in control. Therefore, by the endline, roughly 19% of mothers could not be surveyed. We discuss attrition in detail in section 3.6.

3.3.2 Outcomes

The outcome indices, excluding children's height, were created by combining survey questions. The process involved: (i) transforming each answer into an indicator, where a response on a 5-point Likert scale was coded as 1 for the highest two points and 0 for the remaining three; (ii) aggregating the indicators into a scale; (iii) subtracting the mean of the control group from each scale and dividing the result by the standard deviation (SD) of the control group. We define our outcomes below:

Mental health outcomes. To measure *psychological trauma*, we combined post-traumatic stress disorder and acute stress disorder symptoms (such as distressing memories, avoidance, negative mood, being easily startled, emotional outbursts, etc.) using the simplified Kessler Psychological Distress Scale (Andrews & Slade, 2001) and other survey questions based on the diagnostic criteria laid out by the American Psychiatric Association. To measure *depression*, we used the Center for Epidemiologic Studies Depression Scale-20 (Radloff, 1977). We measured children's psychological trauma and depressive symptoms through an adverse life experience survey (Dyregrov et al., 2000; Neugebauer et al., 2009), which were answered by mothers.

Subjective well-being of mothers. We measure mothers' *happiness*, hope and *aspirations* about the future, and their sense of *belongingness*. As refugees go through the psychological stress of searching for identity (Kumsa, 2006), measures of belongingness inform us about well-being related to their general social identity.

Child development. We measure different developmental progresses associated with their socioemotional, cognitive, and physical development using the Ages and Stages Questionnaire (ASQ-3) questionnaire (Squires & Bricker, 2009). These are *communication*, *gross-motor*, *fine-motor*, *problem-solving*, and *personal-social skills*. Questions are grouped into categories dedicated to assessing a specific set of skills and are also age-specific, e.g., different questions for 2, 4, 6, etc., months-old children. All survey questions were answered by the mothers.

Stunting among children. We explore children's *stunting* by looking at their *height-for-age* z-scores (HAZ). According to WHO (2009), the criterion for stunting is when $HAZ < -2$ SD (i.e. 2 SD below the median in reference population), and severe stunting is when $HAZ < -3$ SD. We use HAZ and dummy variables constructed using these cut-offs as outcomes.

Pre-registration. The outcomes listed above, as well as the survey questions used to calculate them, were pre-registered at the AEA RCT Registry (AEARCTR-0004516). Three additional outcomes (weight-for-height (WHZ) and weight-for-age (WAZ) z-scores, and mother-child relationship) were also pre-registered but dropped due to Covid-19 limitations. Anthropometric enumerators measured the weight of children subjectively (by asking mothers to weigh a 1kg rice sack and then make a 'best guess' of their children's weight), over mobile phones. This data was rather noisy and could not be validated by block-*majhees* later, which is why we have dropped WAZ and WHZ from the paper. Mother-child relationship was measured at baseline but also dropped to reduce the questionnaire length and interview time.

3.4 Sample characteristics and balance checks

We report the balance on observables at baseline between treatment and control groups in Table 1 and the balance on baseline outcomes in Table A1. To derive p-values on tests of equality of means across arms, we regress the variable of interest on the binary treatment with camp fixed effects and standard errors clustered at the unit of randomization. We find our mother and child samples to be well balanced across individual and household characteristics, and average differences in almost all observables are very small. For outcomes measured at baseline (Table A1), again our samples are well balanced. Comparing the differences in distributions of mental health at baseline (shown in Figure A3), we find that the two distributions are statistically similar using a Kolmogorov-Smirnov test (all $p > 0.10$).

Note that we did 33 independent tests—corrections for multiple-hypothesis testing substantially reduce the significance threshold, and, thus, the two significant differences that we observe disappear following such adjustments. In addition, we also compute the normalized differences in means for all variables to show the scale-free differences (Imbens & Wooldridge, 2009).[8] Increasing the sample can also increase the t-statistic, but it does not systematically affect the normalized difference. We find that, out of the 33 normalized differences, 32 differences are lower than $1/8^{th}$ of the combined sample variation and only one difference is below $1/3^{rd}$ (variable 'child victim of at least one camp abuse'). 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 also control for all characteristics that differ in terms of mean or normalized differences when estimating treatment effects.

3.5 Program take-up and session attendance

The initial acceptance rate of the program was 95% (3,499 out of 3,700 invited). About 5% declined due to caring for the elderly or needing permission from their spouses. The weekly participation of the 95% who enrolled was recorded through MV, but only for the treatment group as there were no MVs involved in the control group social gatherings to record attendance.

Out of 1,911 treatment participants, 11 (or 0.6%) participants never attended any sessions, while the remaining participants attended at least one out of 44 sessions.[9] Therefore, the actual take-up among those enrolled was over 99%. Participants attended an average of 20.4 sessions (median is 20) as shown in Figure A2. High participation and attendance were likely due to the restrictions on refugees leaving the camps and the delivery of the program by the trusted organization BRAC. Sessions were also organized within familiar neighborhood settings, delivered by a female neighbor whom participants possibly trust and are familiar with.

[8] For each variable, we first take the difference in means (treatment mean minus control mean) and then divide this difference by the square root of the sum of the variances.

[9] Only five participants' attendance is missing, as we could not match their names in the attendance register to their names in the initial enrolment sheet. Thus, we have an attendance record of 1,906 out of 1,911 in the treatment group. If we consider these 5 participants as 'never-attended', then the total number of participants that never attended any sessions is 16 (or 0.8% of 1,911).

3.6 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). A large portion of the attrition of 19% (or 654 mothers) can possibly be explained by Covid-19. Moreover, the control group had a 14% higher attrition rate than the treatment group. Table A2 showed that the baseline characteristics of mothers/children who attrited were similar to those who remained, with the exception of mothers being the household head ($p < 0.05$). Mothers who dropped out were also 'marginally' newer to the camp and children were slightly shorter, but these differences were small (both $p < 0.10$). We also regress being attrited (equals to 1 if attrited at endline and 0 otherwise) on the treatment indicator, baseline characteristics, and the interaction between the two (Table A3). A joint F-test on the interactions yields a p-value of 0.19, suggesting attrition was not differential by baseline characteristics.[10]

Although we find that observable characteristics of those who attrited versus those who did not are fairly similar across treatment arms, the 14 pp gap in endline participation between treatment and control raises the concern that attrition may bias the treatment effects estimated later in section 4. We address this concern using four different approaches in Appendix B: inverse probability weighting, Lee (2009) bounds, imputing missing data following Kling et al. (2007) and Karlan & Valdivia (2011), and using Horowitz & Manski (2000) bounds. Our main conclusions remain robust using all four approaches. As the 14 pp difference is large, we cannot fully rule out selective attrition based on unobservables. This should be noted as a potential limitation.

As mentioned in section 3.1, about 87% of mothers (or one of their household members) in the study owned a mobile phone, and mobile phone ownership were similar across treatment arms (T-test: $p = 0.916$). The remaining participants were lent mobile phones owned by *majhees* and camp managers. Out of 654 mothers that attrited at endline, 75 did not own a phone (offered but they did not borrow it) and 579 had a phone but did not participate in the endline survey. Among those who participated in the endline survey (2,845 mothers), 381 did not own a phone but borrowed one to participate. In Table A4, we show that baseline characteristics marginally explain phone ownership within the treatment group (joint $p = 0.053$, column 1), but not in the control group (joint $p = 0.290$, column 2) and these characteristics do not jointly differ across treatment arms ($p = 0.60$, column 3).

3.7 Empirical strategy

Treatment effects. To test the impact of the program on mothers and children outcomes, 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. $\boldsymbol{X}_{ijc}$ is a vector of pre-specified controls, measured at the baseline (listed under Table 2). Our results do

[10] Surprisingly, in the treatment arm, average session attendance of those who attritted is 22 compared to 20 among those that did not attrit ($p < 0.01$).

not change if we select controls using the post-double-selection LASSO procedure instead (Belloni et al., 2014). 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. Since trauma and depression indices are based on 'negative' feelings, negative $\hat{\beta}_1$ corresponds to an improvement in mental health. For the remaining outcomes, positive coefficients correspond to more favorable outcomes. We estimate equation 1 using OLS, where β_1 is the intent-to-treat (ITT) effect. As session attendance is very high with over 99% of participants attending at least one session (i.e., taken up the program) and less than 1% not attending any session, we can also interpret the ITT effects as the treatment-on-treated (TOT) effects.

Inference. First, we cluster standard errors at the unit of our randomization. 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.[11] 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. We correct p-values for each outcome that we test 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). Though we do not report FWER p-values using Westfall & Young (1993) in the tables, our conclusions are largely consistent using both methods. Moreover, we aggregate the mental health, mothers' subjective well-being, and child development outcome measures into composite indices to reduce the number of tests (also reported in the main table). Our results are also robust to this adjustment.

4 Main results

4.1 Impacts on mothers

Mental health. Figure 2 and Table 2 report the impact of the intervention on mothers' mental health (Panel A1) and subjective well-being outcomes, such as happiness, aspirations, and belongingness (Panel A2). Column 1 reports treatment effects without controlling for any covariates 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 the mental health of Rohingya mothers. Specifically, mothers that received the treatment experienced a 0.23 SD reduction in psychological trauma ($p < 0.01$) and 0.14 SD reduction in depression severity ($p < 0.01$) relative to mothers in the control group that did not receive the psychosocial program (Panel A1, Table 2). Among mothers who were traumatized or depressed at baseline, the reduction in trauma was slightly higher at 0.26

[11] 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.

SD, and the reduction in depression was substantial at 0.29 SD (a twofold improvement) relative to depressed mothers in the control group. As over 99% of mothers attended one or more sessions, we believe these ITT effects ≈ TOT effects. In Figure A4, we show a correlation between the number of sessions attended and the mental health of mothers in the treatment arm (note attendance was only recorded in the treatment arm), with negative linear fits indicating that higher attendance is correlated with better mental health (pairwise correlation tests: $p < 0.01$ in Plot A and $p = 0.01$ in Plot B).

In comparison to the short-run impacts of other mental health interventions in developing countries (Rahman et al., 2008; Patel et al., 2017), our estimated impacts are relatively smaller. One potential explanation is that the control group in our study also engaged in weekly social gatherings, which is not often seen in previous studies. Social interaction has been shown to have a positive effect on mental health (Nezlek et al., 1994), so the well-being of the control group participants may have improved, reducing the size of the impact. Nonetheless, even with the presence of a placebo, our findings still match the effect sizes seen in a recent non-therapeutic study among Rohingya refugees in Bangladesh (Hussam et al., 2022). We consider these effects to be lower bounds because the COVID-19 pandemic unexpectedly disrupted the intervention halfway through. This is because only half the sessions followed the original protocol, while the remainder were shortened, delivered by phone on a one-to-one basis, and likely less intense due to reduced dosage (20 minutes vs. 60 minutes).

Other well-beings. We then consider outcomes related to mothers' subjective well-being in terms of happiness, aspirations for the future, and belongingness (Panel A2, Table 2). We find that the happiness and belongingness of mothers in the treatment group increased by 0.12 SD ($p < 0.05$) and 0.18 SD ($p < 0.01$) respectively relative to mothers in the control group. However, in terms of aspirations, the treatment effect is muted. We also illustrate these treatment effect estimates in Figure 2, 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' result, we also present results by child's gender (graph A) and by exposure to violence during the conflict in Myanmar (graph B). We do not find the impacts to vary by these characteristics.

The results on other well-being are important because happiness is a fundamental component of well-being, signifying positive emotions and a positive outlook, which can be especially important for those experiencing hardship in refugee camps. Additionally, a sense of belonging encourages integration into the host community, an important factor for refugees who often grapple with isolation and exclusion. The intervention's focus on addressing present mental health needs likely explains why there was no noticeable effect on participants' future aspirations. Refugees may face distinct barriers to developing long-term aspirations due to the uncertainty of their situation and lack of access to economic resources, so focusing on present mental health needs might not be sufficient to move such outcomes. Interventions targeted at future planning (e.g., goal setting, financial planning, etc.), which we did not provide, might be more appropriate.

Are the mentally unhealthy catching up to the healthy? As an exploratory analysis, 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 health and subjective well-being (Panel A1, Table A5). We find that the treatment group mothers who were depressed at baseline caught up to the depression severity of the control group mothers who were healthy at baseline. This

is evidenced by the statistically insignificant coefficients. On the other hand, the treated mothers who were traumatized at baseline surpassed the control group mothers in terms of trauma by 0.20 SD, which is a significant difference at 1% level.[12] In addition, the treated mothers also had higher subjective well-being than the control group mothers (Panel A2, Table A5). This suggests that the intervention was not only successful in alleviating mental distress but also in improving mental health beyond that of mothers who were considered mentally healthy at baseline.

4.2 Impacts on children

Mental health. Figure 2 and Table 2 also report treatment effects on child outcomes. The results indicate that treated children showed an improvement in their levels of trauma and depression severity (Panel B1). Specifically, the children in the treatment group had a decrease of 0.10 SD in trauma ($p < 0.10$, column 2) and a decrease of 0.12 SD in depression ($p < 0.05$, column 2) compared to those in the control group. When we focus on only those children who were traumatized and depressed at baseline, we found results similar to those found for mothers. The impact of the treatment on trauma for these children was 0.13 SD ($p < 0.10$, column 3), slightly higher than the aggregate impact of 0.10 SD. Meanwhile, the treatment had a two-fold larger impact on depression than the pooled impact, resulting in a 0.24 SD decrease ($p < 0.05$, column 3). Additionally, the children whose mothers received the treatment and had trauma or depression at baseline had better mental health following the intervention compared to those in the control group whose mothers were mentally healthy at baseline (Panel B1, Table A5). The correlation between the number of sessions attended by mothers and the mental health of children in the treatment group was also negative, implying that more attendance was correlated with better mental health outcomes for children (pairwise correlation tests: $p = 0.09$ in Plot C and $p = 0.03$ in Plot D) as observed in Figure A4. Therefore, the findings on the mental health of children closely mirror those of their mothers. This is likely because children spend a significant amount of time with their mothers, leading to a strong transmission of mental health from mothers to children.

Socioemotional, physical, and cognitive development. We then analyzed the impact of the intervention on various skills development in children (Panel B2 in Table 2 and Figure 2). The results showed that the intervention significantly improved the communication 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 ($p < 0.10$) compared to the control group. While the improvements in the first three domains were statistically significant at the 1% level, the impact on social skills was weaker and only marginally significant at the 10% level. In terms of fine-motor skills, we do not find any statistically significant treatment effect. Furthermore, analogous to the results for children's mental health, children in the treatment group whose mothers were mentally unhealthy at baseline had improvements in socioemotional, physical, and cognitive development outcomes that either surpassed or were comparable to those of children in the control group whose mothers were mentally healthy at baseline (Panel B2, Table A5), suggesting a close connection between the

[12] Note that mental health is measured using scales to create depression/trauma scores (where higher score corresponds to poor mental well-being). Here, crossing a certain threshold in the score implies being mentally unhealthy. In the literature, the threshold is $1/4^{th}$ of the aggregated score. For instance, CESD-20 is scored between 0-60, and exceeding 15 implies being depressed. Therefore, it is possible for the mentally unhealthy to surpass the scores of the mentally healthy.

development of children and the mental health of their mothers.

These results can be directly mapped into the intervention's activities. Sessions included free play with age-appropriate toys, promoting children's curiosity, communication, and physical development. Activities like peekaboo and toy hunts were designed to specifically target communication, gross motor skills, and problem-solving abilities. The program's emphasis on psychosocial stimulation, which encourages playing with mothers using household items, likely fostered creativity and contributed to improvements in problem-solving and gross motor skills. The focus on large muscle movements may explain the significant gains in gross motor skills but not fine motor skills. Future iterations of the program could benefit from including activities specifically designed to improve fine motor control, such as drawing, building with small blocks, or manipulating small objects.

Height and stunting. In terms of child malnutrition (Table 3), we evaluate the impact of the intervention on stunting. We used height-for-age z-scores (HAZ) as a measure of malnutrition to focus on children's skeletal growth retardation. Our results showed that the intervention was successful in increasing the HAZ of treated children by 0.52 SD ($p < 0.01$), which translates to 1.58 centimeters (column 3, Panel B). At the lower end of the distribution, we observed that stunting and extreme stunting fell by 7 pp (or 10%) and 13 pp (or 22%), respectively. These results were consistent across gender, as there was no difference between male and female children in terms of the reduction in nutritional deprivation (columns 4-6).

This result aligns with research by Nahar et al. (2009) who showed that even brief psychosocial stimulation interventions (of two weeks) can lead to improvements in growth outcomes (0.39 SD units) among children in Bangladesh. This suggests that psychosocial stimulation may play a crucial role in promoting linear growth in children facing adversity. However, it is important to acknowledge the existence of mixed findings in the literature (Super et al., 1990; Hamadani et al., 2006). Some studies, like Nahar et al. (2009), suggest that the impact might be mediated by changes in mothers' parenting behavior. Psychosocial interventions may empower mothers with knowledge and skills to provide better care for their children, including improved mother-child interactions, hygiene behaviors, and a more stimulating home environment. These changes in maternal behavior can then contribute to improvements in children's anthropometric outcomes. Another component of our intervention might also explain this result: maternal mental health. Rahman et al. (2004) shows a strong relationship between poor maternal mental health and child malnutrition. This is because mental health challenges can negatively impact a mother's ability to provide nurturing and attentive care, a crucial factor in child development. Improving maternal mental health can lead to reductions in stunting and wasting among children by fostering positive mother-child interactions and potentially shifting childcare practices (Rahman et al., 2004).

It is important to note a limitation of our study. Due to COVID-19 restrictions, anthropometric measurements were collected through self-reporting by mothers. Mothers were asked to measure their children's height using their right hand and index finger, with these "hand-finger units" later converted to centimeters following Asadujjaman et al. (2019). While these measurements were partially validated by community leaders (which we discuss in details in section 4.3), this method introduces some potential for error compared to in-person measurements by trained personnel. Despite this limitation, our method of measuring height over the phone could be a valuable tool in research settings where in-person assessments are challenging or impossible. Increased conflict and instability around the world limit social science researchers' access to vulnerable populations, mak-

ing remote data collection methods like ours increasingly important.

4.3 Robustness checks

Social desirability bias. 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 psychosocial support for a year might feel more inclined to provide favorable responses to enumerators relative to control group respondents.

We rule out these alternative mechanisms in several ways. First, we observe precise null effects on children's fine-motor skills and mothers' aspirations for the future, and marginal improvements in children's psychological trauma and personal-social skills, which assuage these concerns to some extent. Second, mothers in the control clusters were also enrolled by BRAC in weekly social gatherings but were not told which intervention arm they were part of (i.e., a placebo). Thus, demand effects should have been present in both treatment arms. Third, even if social desirability bias inflated positive responses in the treatment group, it would not explain the stronger treatment effect observed among participants with higher baseline depression. Fourth, enumerators were also blind to the treatment, and many of the child development questions were validated by enumerators during the interview, such as asking mothers to check and then report on how quickly the child grabs the mother's finger, whether their child follows a toy when moved around, can jump, responds to mother's calling, etc.[13] Fifth, all survey outcomes were measured using widely used and validated scales, such as the Center for Epidemiological Studies-Depression scale, which is less vulnerable to demand effects (Barker et al., 2022). This is because the scale's focus on specific symptoms is less likely to be influenced by intervention participation, compared to more direct questions about mental health status. Symptoms like "I was bothered by things that usually don't bother me," "I had trouble keeping my mind on what I was doing," or "I talked less than usual" might be difficult to intentionally manipulate, especially among refugee women with low education levels. Therefore, we do not expect that our psychoeducation treatment would have significantly taught respondents how to strategically answer mental health questions.

Finally, we closely followed Dhar et al. (2022) to measure respondents' general tendency to provide socially desirable responses using the Marlowe-Crowne scale at baseline. This scale was developed by psychologists and has been validated in various contexts and disciplines. Our results show that the treatment effects were not significantly different for those with high or low social desirability bias (SDB) scores and it remained statistically significant among those with low SDB scores. This robustness check, therefore, suggests that our main results are less likely to be a product of experimenter demand effects. We report this result in Table 4.

Measuring height. A major challenge in this project was collecting anthropometric outcomes during the Covid-19 pandemic, but expert anthropometric enumerators from BRAC Bangladesh successfully collected the heights of children over mobile phones. However, some concerns remain such as measurement accuracy, correlation with mothers' opinions, and potential experimenter demand

[13] The survey was conducted by BRAC Institute of Governance and Development (BIGD), and a team with no connection with the program.

effects.

To assess the accuracy of mothers' reports of children's heights and weights, we analyzed the change in height measurements from baseline to endline. Only 3% of mothers reported a height decrease of 5 millimeters or more, with 20% reporting any decrease, indicating that 80-97% of mothers accurately measured their children's height. In fact, following the survey, block-*majhees* randomly visited participating mothers (during Covid-19) to check how accurately mothers measured their children's height but did not find any notable inconsistency in measurements (also discussed in section 3.3.1). We also found no correlation between mothers' opinions on children's growth and their height measures, suggesting that their opinions did not affect their reporting (see Table A7). Finally, the results of our heterogeneous treatment effects analysis using the Marlowe-Crowne scale showed that the treatment effects on HAZ were statistically significant even among mothers with low SDB scores (Table A8), suggesting that remote measures of height were possibly not subject to experimenter demand effects.

Judgment of mothers. Our next concern is whether the results on child development outcomes, reported by mothers, are influenced by the mothers' ability to judge/pay attention to their child's behavior. This is because mental health can impact a person's attention to detail and short-term memory, which can affect their judgment (Zuckerman et al., 2018; Keller et al., 2019). For example, a depressed mother may not have the mental capacity to carefully observe her child or recall important events that are indicative of child development, whereas a non-depressed mother may not face such issues.

We are confident that our results are not the product of such 'judgment bias'. Firstly, the enumerators from BRAC were instructed to be very patient with our respondents to allow them enough time to recall and answer questions carefully. Secondly, many of the questions about child development were validated by the enumerators during the interview, such as the child's ability to grab the mother's finger, follow a toy when moved, jump, respond to the mother's calling, arrange toys vertically or horizontally, etc. Lastly, to address this concern empirically, we re-estimated the treatment effects reported in Table 2 by excluding mothers who showed improvement or change in their mental health. The assumption is that for mothers who remained mentally the same (i.e., depressed mothers remained depressed and non-depressed mothers remained non-depressed) at the endline, their attention to detail and hence judgment should remain constant. If we observe statistically significant treatment effects on child development outcomes among this sample, then mothers' 'judgment bias' cannot be explaining our findings. The results of these conservative estimates are reported in Table A9, which shows that our main findings on child development outcomes are robust even with such extreme adjustments.

Contamination check. Another challenge during the intervention was the possibility of contamination between the control and treatment arms. We address this concern in the following ways. First, we used a cluster-RCT to randomize treatment at the block (i.e., neighborhood) level. Second, the average distance from the treatment to the control blocks was about 70 meters, with multiple non-intervention blocks in between. However, as treatment assignment was done randomly, there were some control blocks with adjacent treatment blocks and vice versa. This allows us to empirically test whether women in control blocks with adjacent treatment blocks experienced any improvements in their mental health.[14] We test this by carrying out a heterogeneity analysis, where

[14]This data is only available for about 1,800 respondents, as the distance data was collected from this interactive map

we interact the treatment dummy with another dummy that captures whether a block has any adjacent treatment block. This result is reported in columns 1 and 6, Table A10. We do not find any statistically significant evidence for improvement in mental health in control blocks with adjacent treatment blocks or for augmented treatment effects in treatment blocks with adjacent treatment blocks. Next, we repeat this exercise with a categorical 'adjacent' variable (with four categories, between 0 and 3, where 3 corresponds to having 3 adjacent treatment blocks). This result is also statistically insignificant (columns 2 and 7). Using the *proportion* of adjacent treatment blocks to total adjacent blocks also does not change this result (columns 3 and 8).

Finally, instead of adjacent blocks, we use the number of treatment blocks within the 200 and 400 meters radius of each block to check whether having more treatment blocks within 200/400 meters radius improved the mental health of control group women (or augmented the mental health impact among the treatment group women). These results are reported in columns 4-5 and 9-10 in Table A10. We again do not find any evidence of contamination in our camps. One plausible reason is that male household heads are overly protective and conservative in this culture (Beech, 2017; Tayeb, 2021b), which might have discouraged or prevented women from leaving their own blocks that could have caused contamination. Moreover, social distancing rules were implemented after 24 sessions, which also restricted the socialization of women across blocks during the intervention.

4.4 Discussions of potential mediators

Mothers' outcomes. Mothers may have experienced improvements in their mental well-being either *directly* or *indirectly* through mediators. The *direct* channel refers to core activities of the intervention such as conducting mental tasks and sharing emotions (psychoeducation), practicing breathing exercises, and engaging in play activities with other participants and children on a weekly basis. On the other hand, the *indirect* channel pertains to information provided during the *My Well-Being* step in each session, which broadly covers encouraging mothers' personal habits related to physical health, better communication with spouses, seeking help, and maintaining social relationships.

In Panel A of Table 5, we investigate these four potential *indirect* channels. We do not observe any significant effects on these potential mediators, indicating that they are unlikely to be plausible channels. The statistically insignificant treatment effect on intimate partner relationships also suggests that our program did not cause relationship friction or, possibly, initiate intimate partner violence. This is important because the program required women to spend an hour every week at a neighbor's house, but the male household heads in this culture are overly protective and conservative in nature (Tayeb, 2021b). One possible explanation for the lack of adverse outcomes is that the sessions were held in the participants' own neighborhoods, specifically in the home of another Rohingya woman, whom they may be familiar with and trust.

We can also rule out socialization as a potential mechanism. The sociological theory of social ties and mental health suggests that socialization can help improve the psychological well-being of people in emotional hardship ('stress-buffering' mechanism) as well as in distress-free conditions ('main effects' mechanism) (Cohen et al., 2000; Kawachi & Berkman, 2001). As mothers in the con-

of the camps in mid-2021, and many block ID numbers from our dataset could not be matched with that in the map as many block ID numbers have changed since 2017 and the map might be showing the updated ID numbers. Note that BRAC does not have this distance information.

trol group also participated in unstructured social gatherings, this potential factor was kept constant across arms by design and is unlikely to be a mechanism. Thus, in this context, the *direct* channels, such as the psychotherapeutic aspects, are more probable.

Children's outcomes. An underlying theory of change suggests that improving mothers' mental well-being would motivate them to prioritize their children's adequate nutrition (e.g., source more food) (Patel et al., 2004; Rahman et al., 2008). However, due to the constraints of living in refugee camps, there are limited opportunities to act on such intentions. We dismiss this possibility because our subsample analysis in Table A9 suggests that the development of children continues to progress, even if their mother's mental health status remains the same (see 'judgment of mothers' in section 4.3).

There are two possible ways in which the program may have *directly* influenced the development of children. Firstly, the psychosocial stimulation provided during *Baby's growing up* step, which includes mother-child interactions through play activities both during the sessions and at home, is a powerful approach for improving the growth and development of children (Grantham-McGregor et al., 1991), with many long-term economic benefits (Heckman et al., 2013; Gertler et al., 2021). In fact, stunted children that only received psychosocial stimulation (and no nutritional supplements) in Grantham-McGregor et al. (1991) were able to catch up to the level of non-stunted children. Second, children participated in free-play activities with age-appropriate toys during the sessions. According to Goldstein (2012), free-play activities in early childhood can have significant emotional and behavioral benefits, such as reducing fear, anxiety, and stress while increasing resilience, as well as social benefits like enhancing empathy, sharing, attention, and attachment, and physical benefits like developing motor skills, increasing flexibility, balance, and coordination. Therefore, a considerable portion of the treatment effects on children's cognitive, physical, and socioemotional development can be directly attributed to psychosocial stimulation and play activities.

There are also scopes for mediated or *indirect* impacts on child development, mostly related to the parenting advice given in the *Baby's growing up* step of each session. Table 5, Panel B, identifies mothers' time input as a significant mediator ($p < 0.01$), with treated mothers spending an additional 1.5 hours daily with their children relative to control group mothers, as also found by Bhalotra et al. (2020) and Vlassopoulos et al. (2024). Of course, in a non-refugee, non-camp context, this mechanism might be somewhat problematic, as accommodating additional 1.5 hours every day for children might mean sacrificing leisure time or less participation in income-generating activities by mothers. However, such opportunity costs for mothers are very small or close to zero in this context because—as mentioned in section 2—refugees cannot leave their designated camps or be employed outside. As a result, mothers mostly spend idle time in their homes after finishing household chores.

We also find that treated mothers were less likely to allow their children to walk or play barefoot ($p < 0.05$) and exhibited less negative parenting behavior ($p < 0.10$), indicating that improvements in mothers' health behaviors toward their children are other potential mechanisms for children's development. However, there was no gender bias in negative parenting, although mothers tended to be more cautious about their sons walking or playing barefoot than their daughters, which is consistent with son preference in South and Southeast Asian countries (Barcellos et al., 2014; Kabeer et al., 2014). Fathers' time input on children, mothers' breastfeeding time or frequency, seeking help for babysitting, and discouraging fathers from smoking indoors were not found to be potential mediators.

5 Heterogeneous treatment effects using machine learning

To understand who benefited the most versus the least from this program, we use a machine learning method developed by Chernozhukov et al. (2020) to examine the heterogeneity in impacts. First, it splits the sample into two equal parts, 'auxiliary' and 'main' sample. From the 'auxiliary' sample, it then generates proxy predictors, $S(Z)$, using machine learning algorithms (in this case, Random Forest) for the conditional average treatment effect (CATE) denoted by:

$$s_0(Z) = E[Y_1|Z] - E[Y_0|Z] \tag{2}$$

where Z is a vector of covariates, Y_1 is the outcome for participants in the treatment group and Y_0 is for control group. Using $S(Z)$, it then generates predictions for the main sample to extract three important properties of $s_0(Z)$: (i) the best linear predictor or BLP|reports the average treatment effect estimates (ATE) and tests for heterogeneity based on Z using machine learning methods (HET); (ii) GATES|calculates group average treatment effects by dividing participants into quintiles based on the extent of their response to the treatment (i.e., least versus most affected); and, (iii) classification analysis or CLAN|which compares the average characteristics of participants in the least and most affected groups. To economize on space, we report BLP and GATES results in Table A11 and CLAN results in Table 6. Table A11 shows that none of the HET parameters are statistically significant at conventional levels (columns 2 and 5), suggesting machine learning algorithms cannot detect the presence of heterogeneity with respect to our set of covariates. Moreover, the differences between the most and least affected quintile groups are also statistically insignificant (columns 3 and 6 for GATES). Even though there is no heterogeneity by characteristics *jointly*, we are interested in exploring if there is any heterogeneity by baseline characteristics *individually*. This is beneficial for two reasons: it allows us to better understand the treatment effects reported in section 4 (i.e., whether effects are lower-/upper-bounds) and it helps policymakers decide whom to target during scale-up, as it can maximize the benefits of the program at the same implementation cost.

In Table 6, we report CLAN results for the following baseline covariates: baseline trauma and depression, age of mother/child, child's gender, households' exposure to violence in Myanmar, and mothers' exposure to camp-based abuse.[15] The program significantly improved the mental health of mothers, particularly those with poor baseline mental health and high exposure to conflict and abuse ($p < 0.01$). Mothers with less education showed more improvement in trauma ($p < 0.01$), and older mothers showed more improvement in depression ($p < 0.01$). While there was no heterogeneity in children's improvement in mental health regarding psychological trauma, older children with more depressive symptoms at baseline showed more improvement in depression ($p < 0.01$).

The program also had significant positive effects on the development of older children and those from households with less exposure to violence in Myanmar, as well as mothers with less experience of abuse in the camps (all $p < 0.01$).[16] Children with more depressive symptoms at baseline showed marginal improvements in development outcomes ($p < 0.10$). Improvement in stunting was most pronounced among older children and those from households with more exposure

[15] Note that we cannot examine heterogeneity by household income, expenditure, employment status, etc. because Rohingyas cannot work or earn in this context.

[16] When we plot treatment effects over children's age for the disaggregated development outcomes (Figure A5), we find that the treatment-control gap gradually widens with children's age in case of communication, gross-motor, and problem-solving skills.

to violent conflict ($p < 0.01$ and $p < 0.05$, respectively), while other differences were not statistically significant.

In Appendix C, we also examine heterogeneity using the traditional "interaction" approach by interacting the treatment dummy with the covariates and testing if the coefficients on the interactions are statistically significant. These results are less sensitive in capturing heterogeneity compared to the machine learning approach. Findings, such as heterogeneity by exposure to violence, remain robust using the "interaction" approach.

6 Lessons from the field and cost effectiveness

Fieldwork challenges. We faced various challenges during the intervention, which the readers and interested policymakers should pay attention to. First, the session facilitators (MVs) were mostly illiterate and needed to be trained by psychosocial experts using pictorial materials. They also had no prior experience in conducting group sessions, so capacity building was necessary and MVs were supervised as needed.[17] Second, the intervention faced difficulties during the Covid-19 pandemic lockdown, as in-person sessions had to be shifted to mobile phone-delivered sessions. BRAC initially had mobile numbers of only 42% of households, so the remaining phone numbers had to be collected at the onset of the pandemic (by camp managers and block-*majhees*), and households without mobile access were lent phones each week. Additionally, participants often had to reschedule phone-sessions at their convenient time, leading to an increased workload for MVs. However, the workload of 7 mother-child pairs per MV (2.5 hours per week) was believed to not generate additional mental burden for MVs. Hussam et al. (2022), in fact, shows that employment can improve the mental well-being of Rohingya refugees. We, however, did not measure the mental health of MVs during or after the intervention; therefore, our take on the mental cost of MVs is only speculative. Altogether, we were able to effectively address these fieldwork challenges due to BRAC Bangladesh's established research infrastructure and reputation in the refugee camps. BRAC's scale of operations in camps, reputation and trust among refugees were also key factors in attracting participants from conservative backgrounds.

Cost effectiveness. The program has two important characteristics that appeal to policymakers in low- or middle-income countries. First, it is peer-delivered, meaning volunteers (MVs) were utilized and do not require complex or prolonged training. The MVs did not need to be highly educated as session materials were simple and could be presented in pictorial forms. This is crucial in resource-poor environments because trained professionals are scarce and costly, and they often do not have adequate knowledge about the culture or language of the refugees.

Second, is the overall cost of the intervention. The cost of providing mental support to mother-child dyads was $45 on average. 50% of this cost was associated with phone-delivered sessions, including the cost of phone calls, payments for block-*majhees* and camp managers, mobile phones for MVs, and revising session materials (see Table A12 for the breakdown). Without the remote sessions, the cost would reduce to $23 per pair. The cost of setting up and maintaining session locations did not incur because MVs used their homes for the weekly sessions. The $45 cost per mother-child dyad (i.e., $19 for a 0.10 SD reduction in trauma and $31 for a 0.10 SD reduction in

[17] For instance, if MVs had difficulty understanding or recalling pictorial directions, they could contact psychosocial experts for advice and support.

depressive symptoms) is still low compared to other studies.

7 Concluding remarks

We demonstrate in this study through a randomized field experiment on Rohingya refugees that a low-cost program combining psychoeducation, parenting support, and play activities can be successfully implemented in resource-poor settings, such as refugee camps in developing countries. We find improvements in the mental health of both mothers and children. Additionally, the program reduced the prevalence of stunting and severe stunting among children and improved their socioemotional, physical, cognitive, and nutritional development. At a cost of $45 per mother-child pair during 44 weekly sessions over a year (about $1 per session) and availing human resources from the refugee community, this program is very scalable and attractive to policymakers.

One-third of forcibly displaced people are refugees, and the majority of them, approximately 85%, are hosted by developing countries where poverty, hunger, and malnutrition are widespread (UNHCR, 2021b). However, hosting refugees comes at a significant cost to LMICs (Taylor et al., 2016). As a result, governments in developing countries must rely on foreign emergency aid and donations to support refugees. In this situation, limited resources may force host countries like Bangladesh to prioritize local welfare policies over refugee aid, which could potentially harm the human capital accumulation of child refugees and result in negative economic consequences for them later in life. Therefore, our program, which is low-cost and easy to scale, can offer an effective but partial solution to promote the health and well-being of both refugee children and their primary-caregiver mothers. In fact, the home-based HPL program is already being scaled up in Bangladesh by BRAC and benefits over 40,000 mother-child dyads. Due to the pressing humanitarian situation, even the control group is now receiving psychosocial support as part of the expansion.

As the world is currently experiencing multiple large conflicts and millions, mostly women and children, have been forced out of their countries, our findings could also offer important insights into the immediate psychosocial needs of these vulnerable refugees during their resettlement.

References

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

Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *Quarterly Journal of Economics*, *136*(4), 2147–2194.

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.

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.

Attanasio, O., Baker-Henningham, H., Bernal, R., Meghir, C., Pineda, D., & Rubio-Codina, M. (2022). Early stimulation and nutrition: The impacts of a scalable intervention. *Journal of European Economic Association (forthcoming)*.

Barcellos, S. H., Carvalho, L. S., & Lleras-Muney, A. (2014). Child gender and parental investments in India: Are boys and girls treated differently? *American Economic Journal: Applied Economics*, *6*(1), 157–89.

Barker, N., Bryan, G. T., Karlan, D., Ofori-Atta, A., & Udry, C. R. (2022). Cognitive Behavioral Therapy Among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress. *Forthcoming at American Economic Review: Insights*.

Baumgartner, J. N., Ali, M., Gallis, J. A., Lillie, M., Owusu, R., Abubakr-Bibilazu, S., Adam, H., Aborigo, R., McEwan, E., Zhou, Y., et al. (2021). Effect of a lay counselor-delivered integrated maternal mental health and early childhood development group-based intervention in northern ghana: a cluster-randomized controlled trial. *Global Mental Health*, *8*.

Bayer, Y. M. & Osher, Y. (2018). Time preference, executive functions, and ego-depletion: An exploratory study. *Journal of Neuroscience, Psychology, and Economics*, *11*(3), 127.

Beech, H. (2017). 'I'm Struggling to Survive': For Rohingya Women, Abuse Continues in Camps. https://www.nytimes.com/2017/12/23/world/asia/rohingya-women-abuse-myanmar.html. Online; Accessed September 12, 2020.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, *81*(2), 608–650.

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

Bhalotra, S., Baranov, V., 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.

Bhalotra, S. R., Delavande, A., Gilabert, P., & Maselko, J. (2020). Maternal investments in children: The role of expected effort and returns. *IZA Discussion Paper NO. 13056*.

Bhat, B., de Quidt, J., Haushofer, J., Patel, V. H., Rao, G., Schilbach, F., & Vautrey, P.-L. P. (2022). The long-run effects of psychotherapy on depression, beliefs, and economic outcomes. *NBER working paper*.

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., Feigl, A. B., Gaziano, T., Hamandi, A., Mowafi, M., et al. (2012). The global economic burden of noncommunicable diseases. Technical report, Program on the Global Demography of Aging.

Bütikofer, A., Ginja, R., Karbownik, K., & Landaud, F. (2024). (breaking) intergenerational transmission of mental health. *Journal of Human Resources*, *59*(7), S108–S151.

Carneiro, P., Kraftman, L., Mason, G., Moore, L., Rasul, I., & Scott, M. (2021). The impacts of a multifaceted prenatal intervention on human capital accumulation in early life. *American Economic Review*, *111*(8), 2506–49.

Carneiro, P. M., Galasso, E., Lopez Garcia, I. X., Bedregal, P., & Cordero, M. (2023). Parental beliefs, investments, and child development: Evidence from a large-scale experiment. *Journal of Political Economy (forthcoming)*.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2020). Generic machine learning

inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *NBER working paper no. 24678*.

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.

Cohen, S., Underwood, L. G., & Gottlieb, B. H. (2000). *Social support measurement and intervention: A guide for health and social scientists*. Oxford University Press.

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), 27–30.

Cuijpers, P., Muñoz, R. F., Clarke, G. N., & Lewinsohn, P. M. (2009). Psychoeducational treatment and prevention of depression: the "coping with depression" course thirty years later. *Clinical Psychology Review*, *29*(5), 449–458.

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.

Dhar, D., Jain, T., & Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review*, *112*(3), 899–927.

Donker, T., Griffiths, K. M., Cuijpers, P., & Christensen, H. (2009). Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. *BMC Medicine*, *7*(1), 1–9.

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.

Ford, L. (2018). Young lives hang by a thread as past haunts Rohingya mothers. https://www.theguardian.com/global-development/2018/mar/29/young-lives-hang-by-thread-as-past-haunts-rohingya-mothers-bangladesh. Online; Accessed March 11, 2022.

Fortify Rights (2020). The Right to Mental Health for Rohingya Survivors of Genocide in Myanmar and Bangladesh. https://www.fortifyrights.org/mya-inv-rep-2020-12-10/. Online; Accessed Janaury 04, 2022.

Fusar-Poli, P., Correll, C. U., Arango, C., Berk, M., Patel, V., & Ioannidis, J. P. (2021). Preventive psychiatry: a blueprint for improving the mental health of young people. *World Psychiatry*, *20*(2), 200–221.

Gertler, P., Heckman, J. J., Pinto, R., Chang, S. M., Grantham-McGregor, S., Vermeersch, C., Walker, S., & Wright, A. (2021). Effect of the jamaica early childhood stimulation intervention on labor market outcomes at age 31. *NBER working paper*.

Goldstein, J. (2012). *Play in children's development, health, and well-being*. TIE Brussels.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., & Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the jamaican study. *The Lancet*, *338*(8758), 1–5.

Habib, M., Jubb, C., Ahmad, S., Rahman, M., & Pallard, H. (2018). *Forced migration of Rohingya: an untold experience*. Ontario International Development Agency: Ontario, Canada.

Hamadani, J. D., Huda, S. N., Khatun, F., & Grantham-McGregor, S. M. (2006). Psychosocial stimulation improves the development of undernourished children in rural bangladesh. *The Journal of Nutrition*, *136*(10), 2645–2652.

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., 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.

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. (2022). The psychosocial value of employment: Evidence from a refugee camp. *American Economic Review*, *112*(11), 3694–3724.

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

Kabeer, N., Huq, L., & Mahmud, S. (2014). Diverging stories of "missing women" in South Asia: Is son preference weakening in Bangladesh? *Feminist Economics*, *20*(4), 138–163.

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.

Kawachi, I. & Berkman, L. F. (2001). Social ties and mental health. *Journal of Urban Health*, *78*(3), 458–467.

Keller, A. S., Leikauf, J. E., Holt-Gosselin, B., Staveland, B. R., & Williams, L. M. (2019). Paying attention to attention in depression. *Translational Psychiatry*, *9*(1), 1–12.

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.

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.

Miguel, E. & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217.

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.

Nahar, B., Hamadani, J. D., Ahmed, T., Tofail, F., Rahman, A., Huda, S., & Grantham-McGregor, S. (2009). Effects of psychosocial stimulation on growth and development of severely malnourished children in a nutrition unit in bangladesh. *European Journal of Clinical Nutrition*, *63*(6), 725–731.

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.

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.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., et al. (2017). The healthy activity program (hap), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in india: a randomised controlled trial. *The Lancet*, *389*(10065), 176–185.

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., Iqbal, Z., Bunn, J., Lovel, H., & Harrington, R. (2004). Impact of maternal depression on infant nutritional status and illness: a cohort study. *Archives of General Psychiatry*, *61*(9), 946–952.

Rahman, A., Khan, M. N., Hamdani, S. U., Chiumento, A., Akhtar, P., Nazir, H., Nisar, A., Masood, A., Din, I. U., Khan, N. A., et al. (2019). Effectiveness of a brief group psychological intervention for women in a post-conflict setting in pakistan: a single-blind, cluster, randomised controlled trial. *The Lancet*, *393*(10182), 1733–1744.

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.

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

Save the Children (2018). Nearly 40 percent of Rohingya children in Cox's Bazar are stunted: new study. https://www.savethechildren.org/us/about-us/media-and-news/2018-press-releases/children-stunted-coxs-bazaar. Online; Accessed October 01, 2020.

Save the Children (2019). Over 80,000 refugee children in Cox's Bazar suffering from severe mental health issues. https://www.savethechildren.org.uk/news/media-centre/press-releases/80000-refugee-children-coxsbazar-mental-health-issues. Online; Accessed March 11, 2022.

Shishir, N. N. (2022). Rising polygamy: Cost of being a woman in rohingya camps. The Daily Star. Accessed on October 01, 2022: https://www.thedailystar.net/opinion/views/news/rising-polygamy-cost-being-woman-rohingya-camps-3086101.

Singla, D. R., Kumbakumba, E., & Aboud, F. E. (2015). Effects of a parenting intervention to address maternal psychological wellbeing and child development and growth in rural uganda: a community-based, cluster-randomised trial. *The Lancet Global Health*, *3*(8), e458–e469.

Song, S. & Teichholtz, S. (2019). Mental health facts on refugees, asylumseekers, & survivors of forced displacement. *American Psychiatric Association*, 1–3.

Squires, J. & Bricker, D. (2009). *Ages & Stages Questionnaires: A Parent-Completed Child Monitoring System*. 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.

Stewart, A. L. & Ware, J. E. (1992). *Measuring functioning and well-being: the medical outcomes study approach*. Duke University Press.

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

Super, C. M., Herrera, M. G., & Mora, J. O. (1990). Long-term effects of food supplementation and psychosocial intervention on the physical growth of colombian infants at risk of malnutrition. *Child Development*, *61*(1), 29–49.

Tayeb, T. (2021a). Four years on, no end in sight to the plight of the Rohingya. https://www.thedailystar.net/opinion/closer-look/news/four-years-no-end-sight-the-plight-the-rohingya-2159946. Online; Accessed Janaury 05, 2022.

Tayeb, T. (2021b). Why are Rohingya women and girls so unsafe in refugee camps? https://www.thedailystar.net/opinion/closer-look/news/why-are-rohingya-women-and-girls-so-unsafe-refugee-camps-2911316. Online; Accessed Janaury 04, 2022.

Taylor, J. E., Filipski, M. J., Alloush, M., Gupta, A., Valdes, R. I. R., & Gonzalez-Estrada, E. (2016). Economic impact of refugees. *Proceedings of the National Academy of Sciences*, *113*(27), 7449–7453.

UNHCR (2021a). Improving Refugee Health Worldwide. https://www.unhcr.org/uk/partners/partners/3fcb53882/improving-refugee-health-worldwide.html. Online; Accessed Janaury 10, 2022.

UNHCR (2021b). Refugee Population Statistics Database. United Nations High Commissioner for Refugees, Geneva. Accessed on February 26, 2022: https://www.unhcr.org/refugee-statistics/.

UNHCR (2022). Global displacement hits another record, capping decade-long rising trend. United Nations High Commissioner for Refugees, Geneva. Accessed on October 02, 2022: https://www.unhcr.org/uk/news/press/2022/6/62a9d2b04/unhcr-global-displacement-hits-record-capping-decade-long-rising-trend.html.

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

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. (2024). Improving women's mental health during a pandemic. *American Economic Journal: Applied Economics*, *16*(2), 422–455.

Walker, S. P., Chang, S. M., Powell, C. A., Simonoff, E., & Grantham-McGregor, S. M. (2006). Effects of psychosocial stimulation and dietary supplementation in early childhood on psychosocial functioning in late adolescence: follow-up of randomised controlled trial. *BMJ*, *333*(7566), 472.

Walker, S. P., Chang, S. M., Wright, A. S., Pinto, R., Heckman, J. J., & Grantham-McGregor, S. M. (2022). Cognitive, psychosocial, and behaviour gains at age 31 years from the jamaica early childhood stimulation trial. *Journal of Child Psychology and Psychiatry*, *63*(6), 626–635.

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

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.

WHO Situation Report (2017). WHO Weekly Situation Report no. 8. https://www.who.int/docs/default-source/searo/bangladesh/bangladesh---rohingya-crisis---pdf-reports/sitreps/2017/weekly-sitrep08-cxbban.pdf?sfvrsn=dc203c83_2. Online; Accessed October 01, 2020.

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

Zuckerman, H., Pan, Z., Park, C., Brietzke, E., Musial, N., Shariq, A. S., Iacobucci, M., Yim, S. J., Lui, L. M., Rong, C., et al. (2018). Recognition and treatment of cognitive dysfunction in major depressive disorder. *Frontiers in Psychiatry*, 655.

Main Figures and Tables

Figure 1: Program timeline

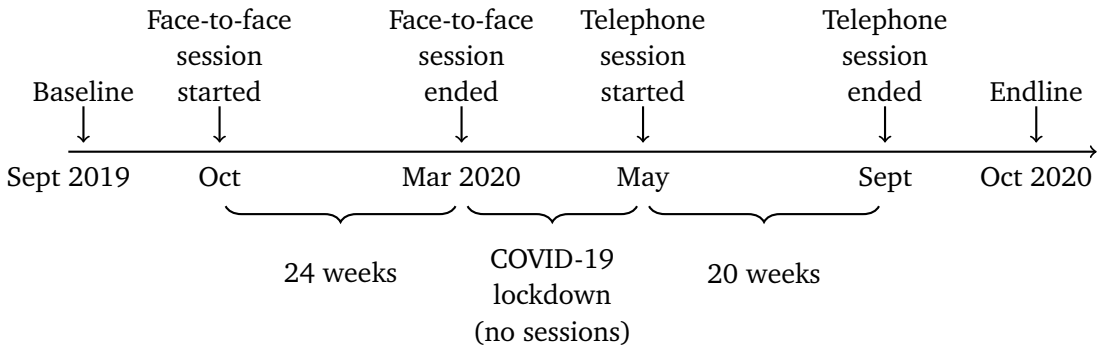
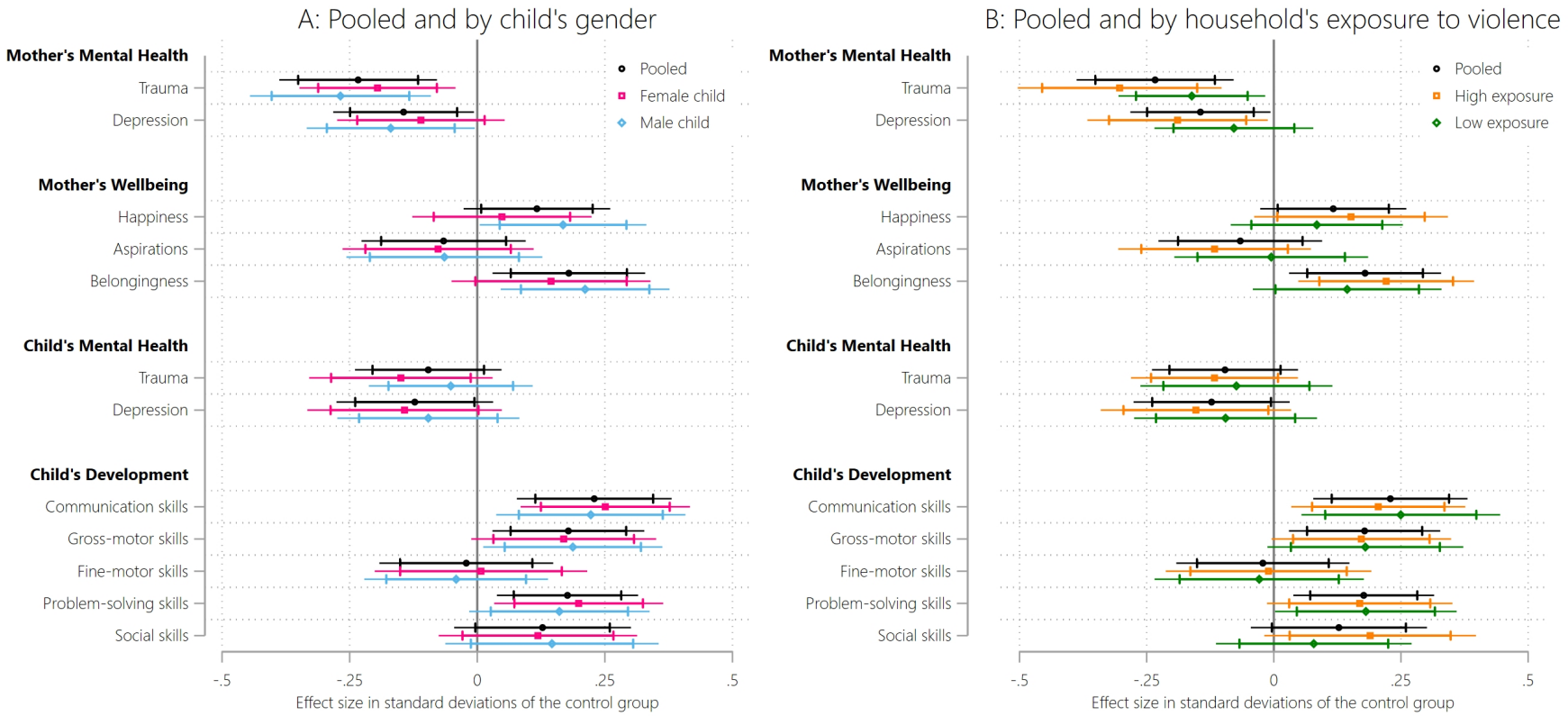


Figure 2: Treatment effects in standard deviations



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

Table 1: 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)			
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). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 2: 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 controlling any covariates. Column (2): treatment effect estimated with full 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 3: Treatment effects on stunting

		Treatment effects				
	Control mean	**Without covariates**	**With all covariates**	**Girl child**	**Boy child**	**Diff (5)-(4)**
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
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)
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.

Table 4: Social desirability bias check

	Mothers' outcomes					Children's 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 2.

Table 5: 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 with all baseline covariates (as in equation 1). Standard errors, clustered at the block level (251 clusters), are in parentheses. Columns (3)-(4) report treatment effects disaggregated by children's gender. Column (5) reports the difference between (4) and (3)—the difference in differences.

Table 6: Heterogeneity Using Random Forest: Classification Analysis (CLAN)

COVARIATES	**Most** (1)	**Least** (2)	**Difference** (3)	COVARIATES	**Most** (4)	**Least** (5)	**Difference** (6)
Outcome: mothers' trauma				**Outcome: children's trauma**			
Mother's trauma at baseline	0.602	0.361	0.232	Child's trauma at baseline	0.493	0.488	0.004
	(0.545,	(0.305,	(0.151,		(0.434,	(0.429,	(-0.078,
	0.659)	0.417)	0.314)		0.552)	0.546)	0.087)
	-	-	[0.000]***		-	-	[1.000]
Age of mother	26.38	25.55	0.815	Age of child	13.95	14.97	0.973
	(25.70,	(24.86,	(-0.142,		(13.19,	(14.23,	(-2.037,
	27.06)	26.27)	1.773)		14.71)	15.72)	0.090)
	-	-	[0.191]		-	-	[0.154]
Attended primary	0.664	0.843	-0.188	Gender of child	0.498	0.522	-0.019
	(0.615,	(0.794,	(-0.257,		(0.439,	(0.464,	(-0.102,
	0.716)	0.892)	-0.119)		0.557)	0.581)	0.064)
	-	-	[0.000]***		-	-	[1.000]
Victimization in Myanmar	0.197	0.100	0.094	Victimization in Myanmar	0.148	0.145	0.002
	(0.182,	(0.086,	(0.073,		(0.134,	(0.130,	(-0.019,
	0.211)	0.115)	0.115)		0.163)	0.160)	0.024)
	-	-	[0.000]***		-	-	[1.000]
Abuse in camp	0.027	0.006	0.022	Abuse in camp	0.013	0.014	0.000
	(0.021,	(0.000,	(0.014,		(0.008,	(0.008,	(-0.007,
	0.033)	0.011)	0.030)		0.019)	0.019)	0.008)
	-	-	[0.000]***		-	-	[1.000]
Outcome: mothers' depression index				**Outcome: children's depression index**			
Mother depressed at baseline	0.457	0.052	0.400	Child depressed at baseline	0.282	0.071	0.216
	(0.412,	(0.007,	(0.336,		(0.240,	(0.028,	(0.158,
	0.502)	0.097)	0.463)		0.324)	0.113)	0.276)
	-	-	[0.000]***		-	-	[0.000]***
Age of mother	26.40	24.90	1.446	Age of child	15.61	13.25	2.435
	(25.73,	(24.17,	(0.505,		(14.86,	(12.51,	(1.346,
	27.07)	25.57)	2.396)		16.37)	13.99)	3.529)
	-	-	[0.006]***		-	-	[0.000]***
Attended primary	0.731	0.719	0.011	Gender of child	0.524	0.496	0.015
	(0.680,	(0.666,	(-0.060,		(0.465,	(0.438,	(-0.068,
	0.783)	0.773)	0.084)		0.583)	0.555)	0.099)
	-	-	[1.000]		-	-	[1.000]
Victimization in Myanmar	0.192	0.098	0.094	Victimization in Myanmar	0.153	0.134	0.015
	(0.177,	(0.083,	(0.074,		(0.138,	(0.120,	(-0.006,
	0.207)	0.113)	0.115)		0.168)	0.150)	0.036)
	-	-	[0.000]***		-	-	[0.313]
Abuse in camp	0.024	0.007	0.018	Abuse in camp	0.018	0.012	0.006
	(0.018,	(0.001,	(0.009,		(0.012,	(0.005,	(-0.001,
	0.030)	0.013)	0.026)		0.023)	0.018)	0.014)
	-	-	[0.000]***		-	-	[0.219]
Outcome: Children's composite development index				**Outcome: Children's stunting**			
Child's trauma at baseline	0.516	0.450	0.073	Child's trauma at baseline	0.484	0.447	0.052
	(0.458,	(0.392,	(-0.009,		(0.426,	(0.388,	(-0.029,
	0.574)	0.508)	0.154)		0.542)	0.505)	0.134)
	-	-	[0.165]		-	-	[0.418]
Child depressed at baseline	0.180	0.118	0.065	Child depressed at baseline	0.153	0.164	-0.015
	(0.139,	(0.076,	(0.005,		(0.110,	(0.122,	(-0.076,
	0.222)	0.161)	0.128)		0.195)	0.207)	0.045)
	-	-	[0.066]*		-	-	[1.000]
Age of child	18.81	10.640	8.073	Age of child	15.97	13.840	2.039
	(18.17,	(10.02,	(7.215,		(15.22,	(13.08,	(1.001,
	19.43)	11.25)	8.931)		16.71)	14.58)	3.102)
	-	-	[0.000]***		-	-	[0.000]***
Gender of child	0.477	0.495	-0.035	Gender of child	0.510	0.492	0.016
	(0.419,	(0.437,	(-0.117,		(0.452,	(0.434,	(-0.067,
	0.535)	0.553)	0.047)		0.568)	0.550)	0.098)
	-	-	[0.803]		-	-	[1.000]
Victimization in Myanmar	0.131	0.167	-0.036	Victimization in Myanmar	0.157	0.135	0.024

	(0.116, 0.146) -	(0.152, 0.181) -	(-0.058, -0.015) [0.002]***		(0.143, 0.171) -	(0.121, 0.150) -	(0.005, 0.045) [0.027]**
Abuse in Camp	0.007 (0.002, 0.012) -	0.021 (0.015, 0.027) -	-0.014 (-0.021, -0.006) [0.001]***	Abuse in Camp	0.016 (0.010, 0.021) -	0.011 (0.005, 0.016) -	0.004 (-0.003, 0.012) [0.449]

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

Note: This table reports CLAN results using Random Forest. 90% confidence interval are in parenthesis; *p*-values for the hypothesis that the parameter is equal to zero are in brackets. 'Most' and 'Least' are the 20% most (top quintile) and 20% least (bottom quintile) affected groups; 'Difference' is the difference in average characteristics between 'Most' and '"Least' affected groups (i.e., most minus least). Outcome of each panel is mentioned at the top. Outcomes that are indices have been control group-standardized. Stunting, Underweight, and Wasting outcomes are dummies where 1 equals stunted, underweight, or wasted growth and 0 otherwise.

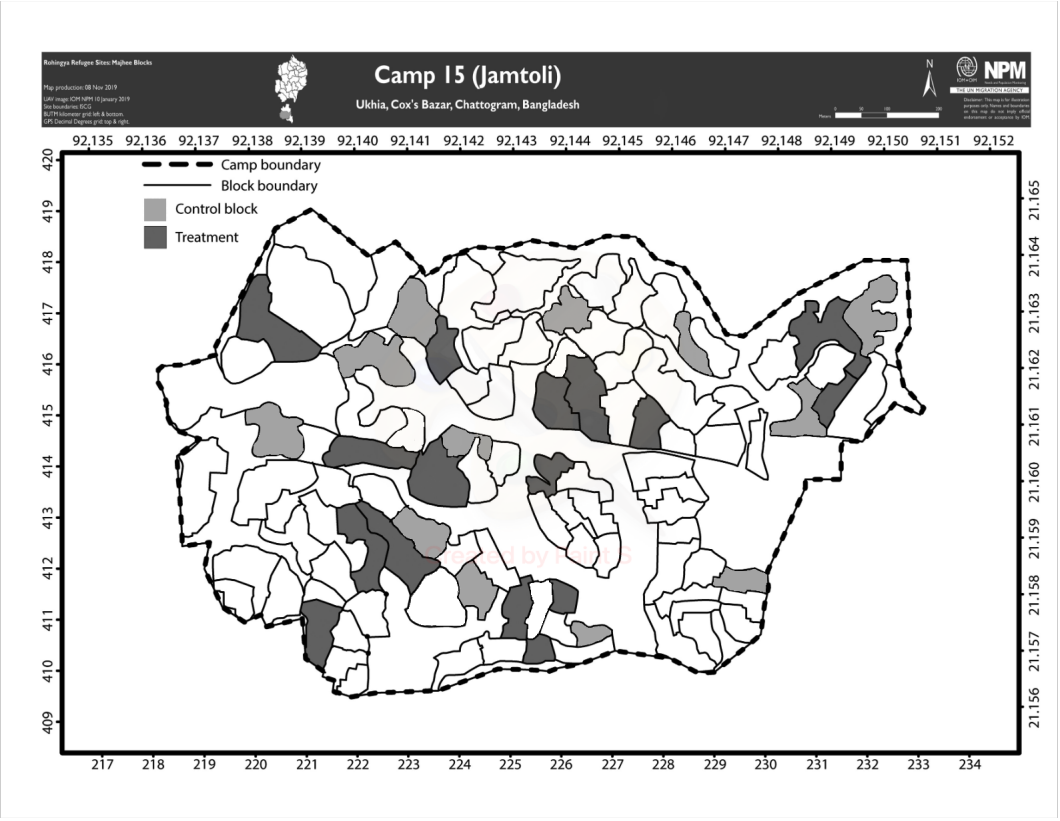
Forced Displacement, Mental Health, and Child Development: Evidence from the Rohingya Refugees

Online Appendix

Asad Islam, Tanvir Ahmed Mozumder, Tabassum Rahman, Tanvir Shatil, Abu Siddique

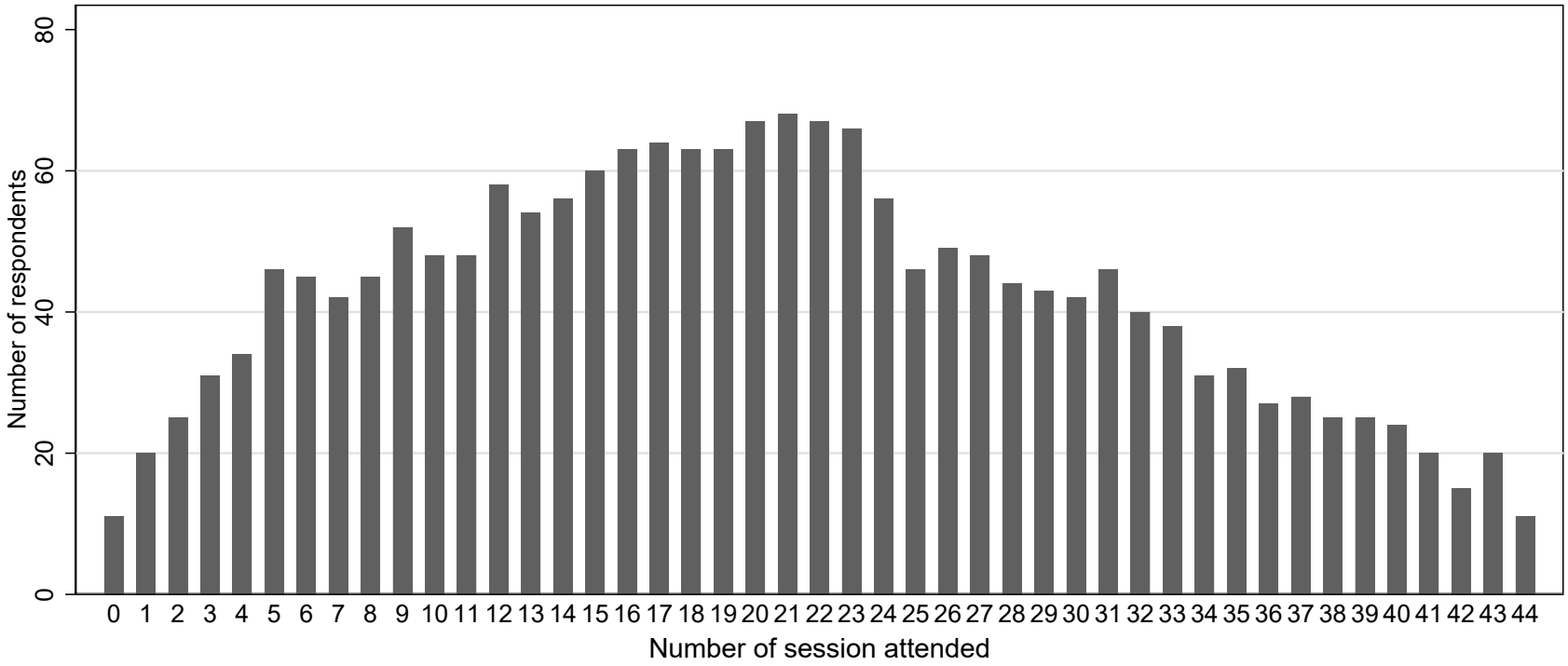
A Appendix: Additional Tables and Figures

Figure A1: Map of a Rohingya camp



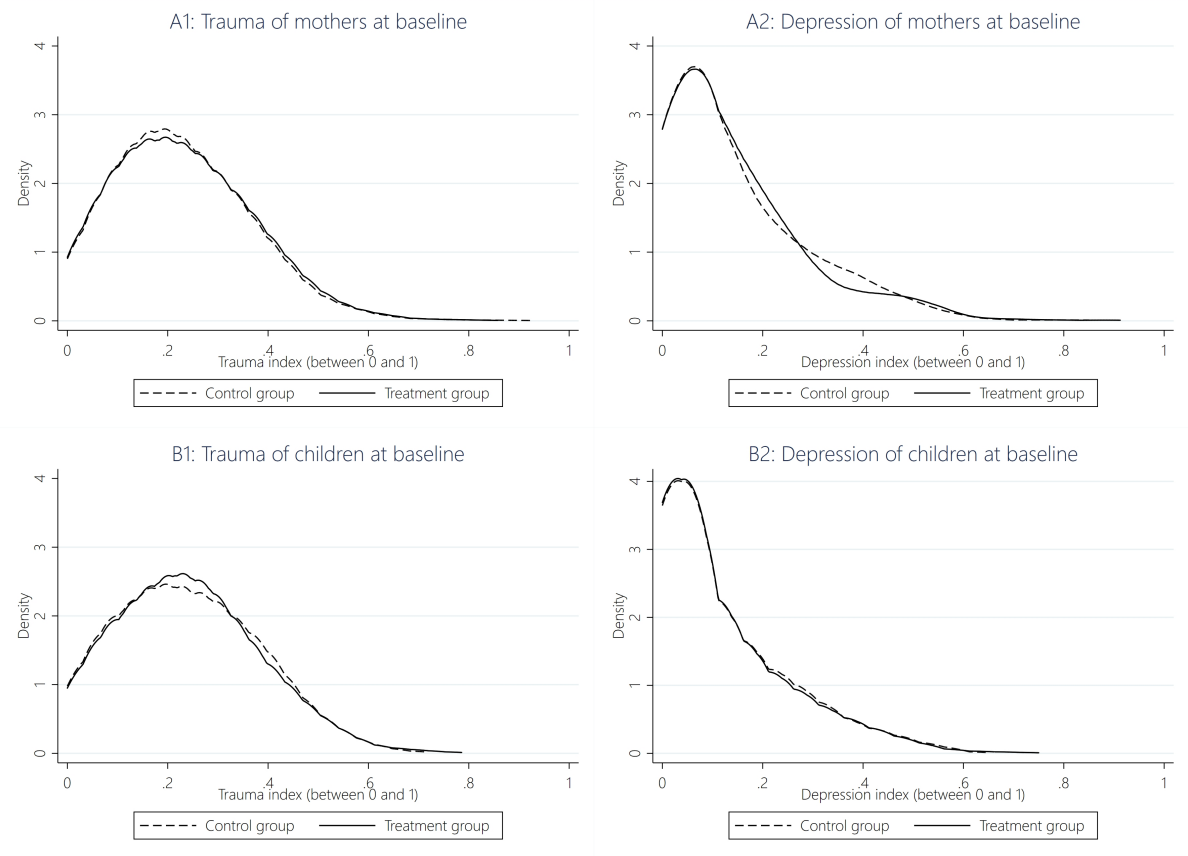
Note: This is a map of Camp 15, showing the treatment and control blocks, and boundaries.

Figure A2: Attendance in treatment sessions



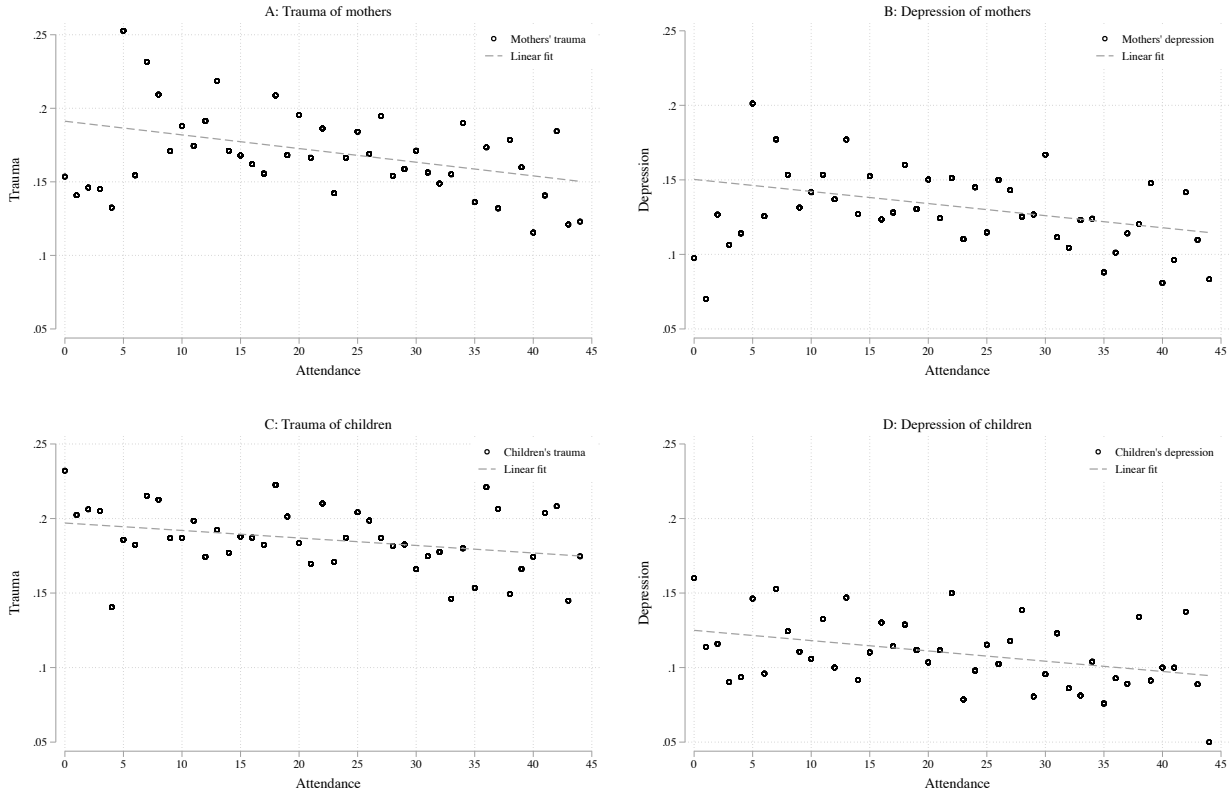
Note: This figure shows the distribution of attendance in treatment group sessions. 0 in the x-axis corresponds to the number of participants that never attended any sessions and 44 corresponds to the number of participants that attended all sessions.

Figure A3: Mental health of mothers and children at baseline



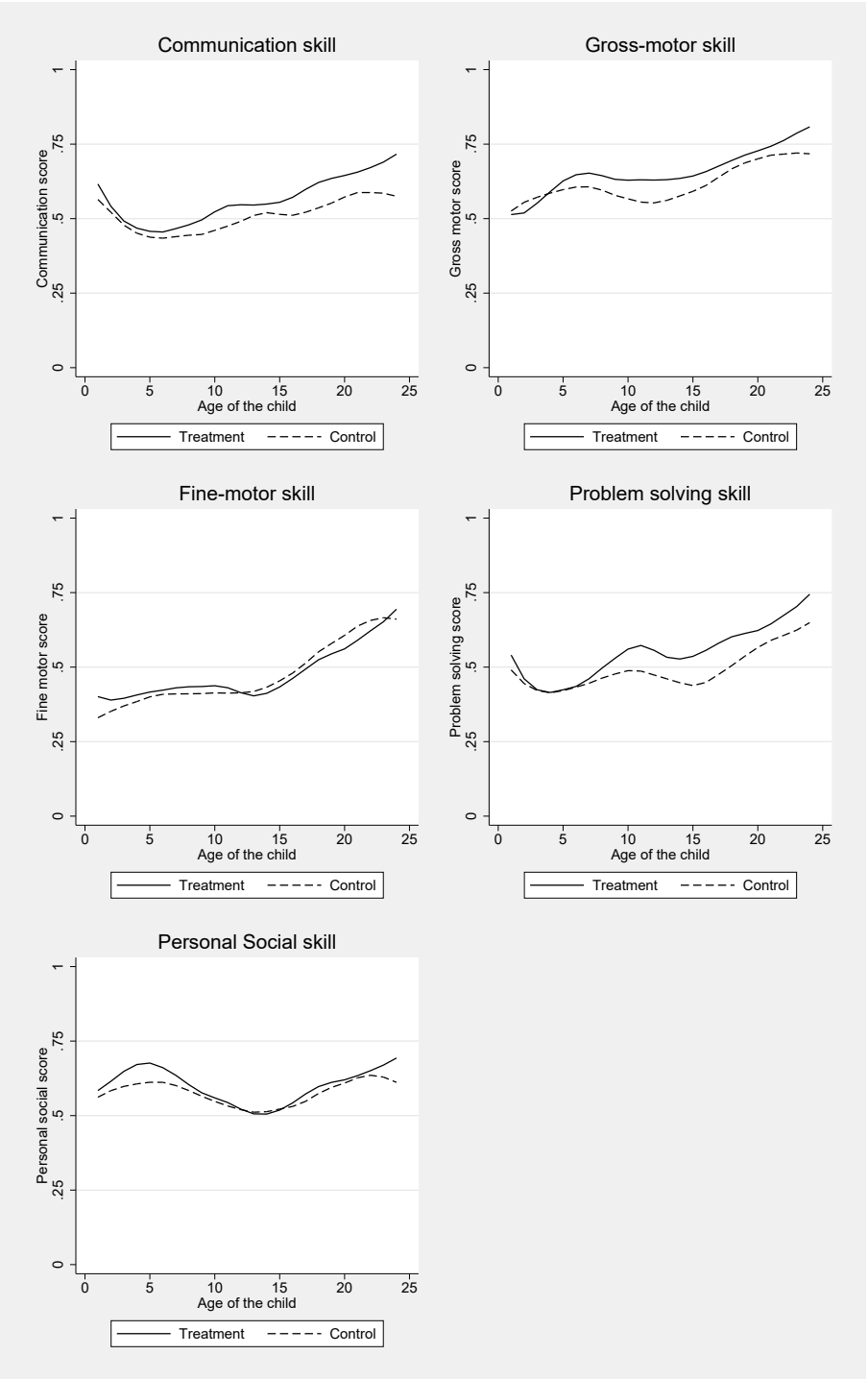
Note: This figure shows the distribution of mental health of mothers (A1 and A2) and children (B1 and B2) at the baseline (estimated from kernel density estimation). Trauma and depression indices are averages of responses to trauma and depression questions, where higher values correspond to more severe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

Figure A4: Correlation between mental health and session attendance



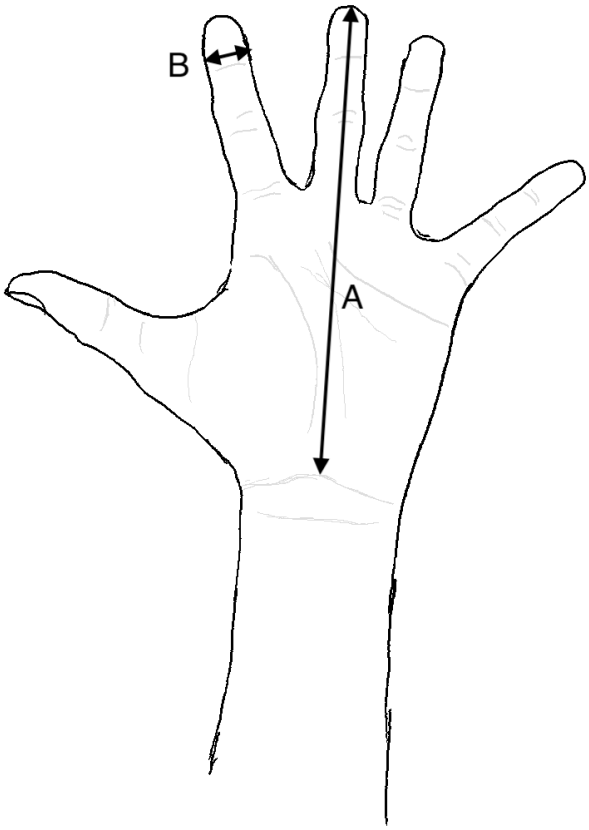
Note: This figure shows the correlation between mental health (y-axis) and session attendance (x-axis). All mental health outcomes have been normalized to be between 0 and 1, where higher value corresponds to poor mental health. Attendance is between 0 and 44, where 44 corresponds to those who attended all 44 sessions.

Figure A5: Treatment effects on skills development, by children's age



Note: This figure shows the treatment effects on skills development by children's age (between 0-24 months). All outcomes are measured at endline.

Figure A6: Over-the-phone measures of length



Note: We used two obsolete anthropic unit of length|hand and finger|to measure children's height over the phone. Here 'hand' length is the distance from the tip of the middle finger to the mid-points of the distal transverse crease of the wrist (i.e., length of A); 'finger' is the width of the index finger (i.e., length of B). All measures were carried out using the right hand, and mothers reported lengths to enumerators in 'hand' and 'finger' units (also, 'half-hand' or 'half-finger' units were considered). Later, using Asadujjaman et al. (2019), we converted these two units into centimeters (cm): hand length= 16 cm and finger= 2 cm.

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)			
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)			

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: 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 A3: 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 abuse)	-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 individual/household characteristics	0.83	0.02	-
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 that we already use. 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 A4: Mobile phone ownership, by treatment

	Treatment	Control	Interaction
VARIABLES	**(1)**	**(2)**	**(3)**
Treatment			0.238
			(0.163)
Age of mothers	-0.001	-0.002	-0.002
	(0.002)	(0.002)	(0.002)
Treatment×Age of mothers			0.000
			(0.003)
Household size	-0.004	-0.003	-0.003
	(0.006)	(0.010)	(0.010)
Treatment×Household size			-0.001
			(0.011)
Mother attended school	-0.012	-0.002	-0.002
	(0.015)	(0.021)	(0.021)
Treatment×Mother attended school			-0.009
			(0.026)
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	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)
Treatment×Duration in the camp			0.000
			(0.000)
Mother receives voucher	-0.017	-0.021	-0.021
	(0.015)	(0.016)	(0.016)
Treatment×Mother receives voucher			0.004
			(0.022)
Husband is alive	0.029	0.036	0.036
	(0.047)	(0.058)	(0.058)
Treatment×Husband is alive			-0.007
			(0.075)
Family member stranded	0.020	0.021	0.021
	(0.022)	(0.029)	(0.029)
Treatment×Family member stranded			-0.001
			(0.036)
HH victimization (conflict)	-0.047	-0.073	-0.073
	(0.070)	(0.063)	(0.063)
Treatment×HH victimization			0.026
			(0.094)
Mothers' victimization (camp abuse)	0.179	-0.024	-0.024
	(0.169)	(0.159)	(0.159)
Treatment×Mothers' victimization			0.203
			(0.232)
Mother is the HH head	-0.001	0.021	0.021
	(0.020)	(0.021)	(0.021)
Treatment×Mother is the HH head			-0.022
			(0.029)
Number of children	0.007	0.013	0.013
	(0.006)	(0.009)	(0.009)
Treatment×Number of children			-0.006
			(0.011)
Age of children	-0.001	-0.006**	-0.006**
	(0.002)	(0.002)	(0.002)
Treatment×Age of children			0.006*
			(0.003)
Gender of children	-0.008	0.000	0.000
	(0.014)	(0.016)	(0.016)

Treatment×Gender of children			-0.008
			(0.022)
Weight of children (kg)	0.013*	0.008	0.008
	(0.007)	(0.007)	(0.007)
Treatment×Weight of children (kg)			0.005
			(0.010)
Height of children (cm)	-0.004**	0.001	0.001
	(0.002)	(0.002)	(0.002)
Treatment×Height of children (cm)			-0.005**
			(0.002)
Child's victimization (camp)	0.147	0.022	0.022
	(0.152)	(0.090)	(0.090)
Treatment×Child's victimization (camp)			0.125
			(0.177)
Observations	1,907	1,586	3,493
R-squared	0.008	0.010	0.009
Mobile ownership	0.8702	0.8690	-
Joint *p*-value on individual/household characteristics	0.053	0.290	-
Joint *p*-value on interactions	-	-	0.603

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 mobile ownership, a dummy variable that equals 1 if a mother (or any household member) has a mobile phone and 0 if she does not. 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 that we already use. All variables were measured at the baseline. Overall mobile phone ownership is roughly 87%.

Table A5: 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 mothers from the treatment arm that were traumatized at baseline (or the mentally unwell) and mothers from the control arm that did not have trauma at baseline (or the mentally well). Similarly, vertical panel Y (depression) includes mothers from the treatment arm that were depressed at baseline (or the mentally unwell) and mothers from the control arm that did not have depression 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 **??**). Covariates are listed under Table 2. For outcomes with ‡, negative coefficients imply more favorable outcomes.

Table A6: 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 2). 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 2.

Table A7: Growth opinions and height measures

	Height↑
VARIABLES	**(1)**
Height (in cm)	0.000
	(0.000)
All Controls	Yes
Camp FE	Yes
Observations	2,840
R-squared	0.025

Robust SE clustered at the block level
*** p<0.01, ** p<0.05, * p<0.1

Note: OLS estimates reported. Dependent variables are mothers' opinions about children's improvement in height (Height↑), which is a dummy variable, where 1 means improved and 0 means did not improve. Independent variable Height (in cm) is the measure of height at endline. Controls are listed under Table 2.

Table A8: Social desirability bias check for HAZ

	HAZ
VARIABLES	**(1)**
Treatment	0.562***
	(0.175)
High SDB	0.383*
	(0.217)
Treatment×High SDB	-0.102
	(0.272)
All Controls	Yes
Camp FE	Yes
Observations	2,840
R-squared	0.106

Robust SE clustered at the block level
*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variables is height-for-age z-score or HAZ. 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 2.

Table A9: Judgment of mothers

	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	0.024	0.025	-0.056	-0.039
	(0.058)	(0.058)	(0.062)	(0.060)
Depression severity	-0.025	-0.030	-0.034	-0.039
	(0.060)	(0.061)	(0.064)	(0.062)
B. Child's development				
Communication skills	0.211**	0.195**	0.218***	0.199***
	(0.085)	(0.081)	(0.072)	(0.070)
Gross-motor skills	0.213**	0.207**	0.158**	0.147**
	(0.084)	(0.081)	(0.074)	(0.071)
Fine-motor skills	-0.037	-0.055	-0.058	-0.073
	(0.094)	(0.087)	(0.082)	(0.077)
Problem-solving skills	0.214**	0.194**	0.167**	0.154**
	(0.082)	(0.078)	(0.069)	(0.066)
Social skills	0.131	0.134	0.081	0.080
	(0.091)	(0.091)	(0.075)	(0.075)
C. Child's height				
Height-for-age z-score	0.500**	0.437**	0.447**	0.362**
	(0.223)	(0.210)	(0.190)	(0.173)
Observations	1,311	1,308	1,893	1,891

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

Note: All panels include women whose mental health remained unchanged from baseline to 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 **??**). Covariates are listed under Table 2. For outcomes with ‡, negative coefficients imply more favorable outcomes.

Table A10: Contamination check

	Mother's trauma					Mother's depression				
	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.

Table A11: Heterogeneity using Random Forest: BLP and GATES results

	(1)	(2)	(3)		(4)	(5)	(6)
	Outcome: mothers' trauma				**Outcome: children's trauma**		
	ATE	**HET**			**ATE**	**HET**	
BLP	-0.226	0.485	-	BLP	-0.113	0.164	-
	(-0.353,-0.107)	(-0.487,1.437)	-		(-0.245,0.019)	(-1.231,1.720)	-
	[0.001]***	[0.654]	-		[0.185]	[1.000]	-
	Most	**Least**	**Difference**		**Most**	**Least**	**Difference**
GATES	-0.365	-0.234	-0.123	GATES	-0.130	-0.096	-0.046
	(-0.624,-0.110)	(-0.457,-0.013)	(-0.469,0.209)		(-0.353,0.104)	(-0.334,0.147)	(-0.353,0.274)
	[0.010]***	[0.077]*	[0.919]		[0.529]	[0.810]	[1.000]
	Outcome: mothers' depression index				**Outcome: children's depression index**		
	ATE	**HET**			**ATE**	**HET**	
BLP	-0.135	0.626	-	BLP	-0.122	0.327	-
	(-0.261,-0.007)	(-0.179,1.462)	-		(-0.261,0.013)	(-0.446,1.135)	-
	[0.077]*	[0.270]	-		[0.155]	[0.894]	-
	Most	**Least**	**Difference**		**Most**	**Least**	**Difference**
GATES	-0.258	-0.069	-0.194	GATES	-0.199	-0.062	-0.143
	(-0.497,-0.019)	(-0.143,0.276)	(-0.116,0.491)		(-0.442,0.038)	(-0.290,0.157)	(-0.458,0.178)
	[0.070]*	[1.000]	[0.453]		[0.208]	[1.000]	[0.729]
	Outcome: Children's composite development index				**Outcome: Children's stunting**		
	ATE	**HET**			**ATE**	**HET**	
BLP	0.196	0.261	-	BLP	-0.070	0.021	-
	(0.042,0.345)	(-0.250,0.708)	-		(-0.120,-0.021)	(-0.296,0.348)	-
	[0.027]**	[0.675]	-		[0.010]***	[1.000]	-
	Most	**Least**	**Difference**		**Most**	**Least**	**Difference**
GATES	0.263	0.095	0.168	GATES	-0.057	-0.056	-0.005
	(0.008,0.507)	(-0.174,0.359)	(-0.182,0.499)		(-0.162,0.040)	(-0.167,0.055)	(-0.148,0.146)
	[0.087]*	[0.984]	[0.675]		[0.487]	[0.641]	[1.000]
BLP	-0.072	0.889	-	BLP	-0.090	0.155	-

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

Note: This table reports BLP and GATES results using Random Forest. 90% confidence interval are in parenthesis; *p*-values for the hypothesis that the parameter is equal to zero are in brackets. ATE is the average treatment effect and HET is the heterogeneity loading parameter. 'Most' and 'Least' are the 20% most (top quintile) and 20% least (bottom quintile) affected groups; 'Difference' is the difference in average characteristics between 'Most' and '"Least' affected groups (i.e., most minus least). Outcome of each panel is mentioned at the top. Outcomes that are indices have been control group-standardized. Stunting is a dummy where 1 equals stunted and 0 otherwise.

Table A12: Program cost

Cost details	**Cost in BDT**	**Cost in USD**
Salary and benefits of Senior Psycho-Social Counselors	432,507	5,088.32
Salary and benefits of Psycho-Social Counselors	1,505,851	17,715.89
Session material development workshop	1,223,543	14,394.62
Hiring, training, and refreshers for mother-volunteers	206,835	2,433.35
Session materials and printing	241,641	2,842.84
Training on play pedagogy for all staff	27,260	320.71
Mobile phone cost and support	3,657,051	43,024.13
Total cost	**7,294,688**	**85,819.86**
Cost per treated mother-child pair (N=1,911)	**3,817.21**	**44.90**

Note: USD 1 = 85 Bangladeshi Taka (BDT), when this study was conducted.

B Appendix: Differential attrition and treatment effects

As highlighted in section 3.6, there is significantly higher attrition in the control group relative to the treatment group ($p < 0.01$). 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. First, we use inverse probability weighting (IPW) to estimate the treatment effects. For this, respondents are weighted by the inverse of their response-probability, which implies that women with characteristics similar to women that are missing at endline are up-weighted in the analysis, whereas those with a high probability to respond at endline are given low weights in the analysis. These attrition-adjusted estimates are almost identical to the unadjusted estimates, which are presented in Table B1 (unadjusted effects in column 1 and IPW-adjusted effects in column 2). Second, following Lee (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. Our conclusions remain largely consistent with Lee (2009) bounds (columns 3-4, Table B1), where most of the treatment effects survive.[1]

Third, following Kling et al. (2007); Karlan & Valdivia (2011), we impute the missing outcome-observations in the treatment arm using the following equation:

$$Missing\ values^{T} = \overline{Y}^{T} + \delta \tag{1}$$

where $\overline{Y}^{T}$ is the mean of mental health outcomes (Y) in the treatment group (T), and $\delta = 0.05$, 0.10, or 0.25 standard deviations. In other words, we first generate the averages of mental health outcome variables in the treatment arm ($\overline{Y}^{T}$) and then create three new variables by adding 0.05, 0.10, and 0.025 standard deviations (δ) to the averages of the outcomes (i.e., $\overline{Y}^{T} + \delta$), respectively. Finally, we impute these newly generated values to the mental health outcomes of attritors (or non-responders) in the treatment group. On the other hand, instead of subtracting 0.05, 0.10, and 0.025 SD to the averages in the control arm, we impute zeros to missing observations in the control arm. This is because, we make these adjustments to control-standardized outcome indices, where the control group has mean 0 already. Since negative values for mental health variables correspond to favorable outcomes, imputing $\overline{Y}^{T} + \delta$ to missings in the treatment arm creates three lower bounds. In contrast, positive values for subjective well-being and child development outcomes correspond to favorable outcomes. Thus, for these outcomes, we impute $\overline{Y}^{T} - \delta$ to missings in the treatment arm and 0 to that in the control arm to generate their lower bounds. Finally, a higher HAZ score is also associated with favorable outcomes, but this z-score is not control group-standardized. Therefore, to create the lower bounds, we impute $\overline{Y}^{T} - \delta$ to missings in the treatment arm and $\overline{Y}^{C} + \delta$ to missings in the control arm, where $\overline{Y}^{C}$ is the mean of the outcome in the control arm (C).

Results using these newly generated lower bounds is presented in Table B2, where columns 2-4 report estimates with $\delta = 0.05$ SD (column 2), $\delta = 0.10$ (column 3), and $\delta = 0.25$ SD (column 4). These three bounds show that our main results would hold even if the outcomes of the attrited sample in the treatment group were 0.25 SD worse on average than that in the control group. In fact, except for mothers' happiness and children's trauma, all other results remain similar to the

[1] Few exceptions are mothers' happiness and children's problem-solving and fine-motor skills.

unadjusted effects (column 1) even for the more extreme $\delta = 0.25$ adjustments (column 4).

Finally, although based on extreme assumptions about attrition, we follow Horowitz & Manski (2000)'s version in Karlan & Valdivia (2011) to create two additional extreme bounds (both lower and upper). For this, we impute on the basis of minimal and maximal possible values to missing information. 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. This gives us the most extreme lower and upper bounds. In a similar manner, instead of imputing minimal and maximal values, we replace 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. This gives us the 2^{nd}-most extreme lower and upper bounds. We report treatment effects using these bounds in columns 5-9 in Table B2. We find that Horowitz & Manski (2000) bounds yield very wide bounds due to imputing extreme values. This is because, this bounds analysis is suitable when outcomes are discrete and attrition is very low (Ozler, 2017). In fact, Karlan & Valdivia (2011) also finds these bounds to be 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 up to 0.25 SD. According to column 1 in Table B2 (same as column 2 in Table 2), the largest effect size for mental health outcomes 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 were much worse-off than attrited mothers in the control group|only changes the effect size by roughly 0.03 SD (from -0.23 to -0.20). This suggests that the mental health of attritors in the treatment arm would have to be extremely poor than non-attritors to change our main conclusions.

Table B1: Treatment effects: Inverse Probability Weighting & Lee bounds

	Treatment effects		Lee (2009) bounds	
	Unadjusted	IPW	Lower	Upper
Dependent variables	**(1)**	**(2)**	**(3)**	**(4)**
A1. Mothers' mental health				
Trauma	-0.233***	-0.234***	-0.470***	-0.160***
	(0.051)	(0.049)	(0.035)	(0.037)
Depression	-0.144***	-0.144***	-0.330***	-0.104***
	(0.054)	(0.052)	(0.030)	(0.034)
A2. Mothers' well-being				
Happiness	0.117**	0.124**	0.011	0.523***
	(0.056)	(0.054)	(0.040)	(0.044)
Aspirations	-0.066	-0.073	-0.295***	0.242***
	(0.062)	(0.061)	(0.040)	(0.046)
Belongingness	0.179***	0.190***	0.076*	0.490***
	(0.057)	(0.055)	(0.043)	(0.044)
B1. Children's mental health				
Trauma	-0.096*	-0.094*	-0.380***	-0.024
	(0.055)	(0.054)	(0.036)	(0.038)
Depression	-0.122**	0.117**	-0.343***	-0.059
	(0.059)	(0.057)	(0.029)	(0.038)
B2. Children's development				
Communication skills	0.229***	0.232***	0.139***	0.609***
	(0.059)	(0.058)	(0.042)	(0.049)
Gross-motor skills	0.179***	0.189***	0.175***	0.482***
	(0.058)	(0.056)	(0.042)	(0.044)
Fine-motor skills	-0.021	-0.017	-0.271***	0.289***
	(0.066)	(0.064)	(0.051)	(0.045)
Problem-solving skills	0.177***	0.172***	-0.027	0.489***
	(0.055)	(0.054)	(0.047)	(0.046)
Social skills	0.128*	0.148**	-0.135***	0.410***
	(0.067)	(0.066)	(0.042)	(0.043)
B3. Children's height				
Height-for-age z-score	0.515***	0.521***	-0.512***	1.487***
	(0.139)	(0.137)	(0.150)	(0.161)

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 2. Column (2) reports the Inverse Probability Weight (IPW) adjusted treatment effects. Columns (3)-(4) report the lower and upper bound treatment effects using Lee (2009) bounds.

Table B2: Treatment effects: Additional bounds analysis

	Unadjusted	Kling et al. (2007) Bounds			Most Extr. Bounds		2^{nd}-Most Extr. Bounds	
	Treatment Effects	$\delta = 0.05$	$\delta = 0.10$	$\delta = 0.25$	Lower	Upper	Lower	Upper
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A1. Mothers' mental health								
Trauma	-0.233***	-0.227***	-0.221***	-0.200***	-1.778***	0.792***	-1.069***	0.212***
	(0.051)	(0.038)	(0.038)	(0.038)	(0.148)	(0.095)	(0.085)	(0.050)
Depression	-0.144***	-0.126***	-0.120***	-0.099**	-2.420***	0.914***	-1.020***	0.233***
	(0.054)	(0.040)	(0.040)	(0.040)	(0.217)	(0.101)	(0.090)	(0.049)
A2. Mothers' well-being								
Happiness	0.117**	0.095**	0.089**	0.068	-0.426***	0.933***	-0.102*	0.579***
	(0.056)	(0.043)	(0.043)	(0.043)	(0.061)	(0.078)	(0.058)	(0.059)
Aspirations	-0.066	-0.054	-0.061	-0.082*	-1.095***	1.091***	-0.864***	0.641***
	(0.062)	(0.049)	(0.049)	(0.049)	(0.087)	(0.098)	(0.080)	(0.068)
Belongingness	0.179***	0.176***	0.169***	0.148***	-0.414***	0.984***	-0.001	0.499***
	(0.057)	(0.046)	(0.046)	(0.046)	(0.063)	(0.081)	(0.057)	(0.058)
B1. Children's mental health								
Trauma	-0.096*	-0.094**	-0.087**	-0.066	-2.089***	0.823***	-1.166***	0.315***
	(0.055)	(0.042)	(0.042)	(0.042)	(0.189)	(0.090)	(0.106)	(0.053)
Depression	-0.122**	-0.111**	-0.105**	-0.084*	-2.574***	0.893***	-1.226***	0.486***
	(0.059)	(0.045)	(0.045)	(0.046)	(0.235)	(0.101)	(0.113)	(0.069)
B2. Children's development								
Communication skills	0.229***	0.226***	0.220***	0.201***	-0.454***	0.867***	-0.398***	0.762***
	(0.059)	(0.048)	(0.048)	(0.048)	(0.069)	(0.070)	(0.068)	(0.063)
Gross-motor skills	0.179***	0.174***	0.168***	0.149***	-0.408***	0.850***	-0.048	0.657***
	(0.058)	(0.046)	(0.046)	(0.046)	(0.062)	(0.072)	(0.058)	(0.068)
Fine-motor skills	-0.021	-0.015	-0.021	-0.040	-0.608***	0.581***	-0.205***	0.403***
	(0.066)	(0.053)	(0.053)	(0.053)	(0.068)	(0.072)	(0.064)	(0.072)
Problem-solving skills	0.177***	0.178***	0.172***	0.153***	-0.400***	0.720***	-0.020	0.577***
	(0.055)	(0.044)	(0.044)	(0.044)	(0.061)	(0.062)	(0.055)	(0.062)
Social skills	0.128*	0.113**	0.107**	0.088*	-0.542***	0.829***	-0.087	0.574***
	(0.067)	(0.052)	(0.052)	(0.052)	(0.072)	(0.075)	(0.068)	(0.069)
B3. Children's height								
Height-for-age z-score	0.515***	0.511***	0.505***	0.486***	-5.898***	5.338***	-2.391***	2.793***
	(0.139)	(0.133)	(0.133)	(0.133)	(0.550)	(0.368)	(0.263)	(0.199)

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 2. Columns 2-4 report treatment effects with moderate bounds following Kling et al. (2007). Columns 5-8 report treatment effects with extreme bounds following Horowitz & Manski (2000) and Karlan & Valdivia (2011). All specifications control for baseline characteristics and standard errors are clustered at the unit of randomization.

C Appendix: Additional discussions and analyses

C.1 Literature in details

Psychotherapy. Psychotherapy, such as cognitive behavioral therapy (CBT), problem management plus (PM+), acceptance and commitment therapy (ACT), psychoeducation, interpersonal psychotherapy (IPT), behavioral activation (BA), etc., are commonly used to mitigate mental health problems. Psychotherapy is often more effective than antidepressants (Cronin et al., 2020), and the downstream consequences of improved mental health include positive behavioral change, human capital accumulation, and better economic decision-making and outcomes in both short- and long-term (Currie & Stabile, 2007; Cuijpers et al., 2016; Singla et al., 2017; Ridley et al., 2020). To better understand our contribution, we have summarized the most relevant studies on mental health interventions in Table C1.

Our intervention includes psychoeducation, which is an established method that combines light-touch psychotherapy and mental-health education (Christensen et al., 2004; Geisner et al., 2006; Reynolds et al., 2017).[2] Cuijpers et al. (2009), a meta-analysis on psychoeducation, shows that psychoeducation treatment can reduce the risk of getting major depression by 38% and can improve depressive symptoms by 0.28 SD. Cuijpers et al. (2009) also finds no evidence of psychoeducation being less effective than other psychotherapy treatments.

Another commonly used psychotherapy is CBT, which helps people change their thinking and behavioral patterns by breaking down problems and reaching solutions. CBT has shown great effectiveness in reducing depressive symptoms among people in low or middle-income countries (LMIC) (Patel et al., 2017; Fuhr et al., 2019; Barker et al., 2022), particularly among mothers of young children (Rahman et al., 2008) and the elderly (McKelway et al., 2022). Positive impacts on depression can also persist in the longer term and subsequently affect the financial empowerment of women and time-input on children (Bhalotra et al., 2020), and several cognitive and non-cognitive skills of the treated (Barker et al., 2022; Bhat et al., 2022).

On the other hand, PM+, which is more light-touch and can be delivered by non-specialists, has also proved to be effective (Bryant et al., 2017), with striking results in reducing depressive symptoms, psychological trauma, and anxiety in post-conflict settings (Rahman et al., 2016, 2019).[3] Moreover, PM+ on the forcibly displaced refugees have been very effective in reducing their depression, trauma, and anxiety symptoms (de Graaff et al., 2020; Acarturk et al., 2022); though, others were less successful (Haushofer et al., 2020). Another variant of CBT is IPT which focuses on solving interpersonal problems and can also be very effective in reducing depression and trauma (Bolton et al., 2003).

Early-childhood stimulation and parenting. Our paper also contributes to the literature on interventions targeting early-childhood development (ECD) (Grantham-McGregor et al., 1991; Campbell & Ramey, 1994; Heckman et al., 2013), particularly to the programs that facilitate psychosocial stimulation through play-activities (Grantham-McGregor et al., 1991; Yousafzai et al.,

[2] Informational and light-touch talk therapy (delivered remotely) have also been proven to be effective in reducing depressive symptoms, stress, and anxiety among people in isolation (Vlassopoulos et al., 2024; Sadish et al., 2021).

[3] In a similar context, Hussam et al. (2022) offered eight-weeks long employment opportunities to Rohingyas living in refugee camps in Bangladesh and finds that the mental benefits from being employed surpass the mental benefits of receiving cash transfers among the refugees. Thus, this study uses a non-psychotherapeutic intervention to address the mental health of adult refugees.

2014, 2016), parenting counseling, or a mixture of the two in LMICs (Singla et al., 2015; Baumgartner et al., 2021). Other variants of the early-childhood psychosocial stimulation program delivered via trained community peers were also found to be effective in improving ECD outcomes (Attanasio et al., 2014, 2022; Amadu et al., 2019).[4]

There is also growing evidence that high-quality ECD interventions and environments can boost human capital accumulation and affect later-life outcomes (Almond & Currie, 2011; Campbell et al., 2014). For instance, follow-ups of Grantham-McGregor et al. (1991)—the influential ECD program in Jamaica that focused on providing psychosocial stimulation and nutrients—show that treated children had improved IQ, educational attainment, earnings, and mental health during adulthood (Walker et al., 2011; Gertler et al., 2014, 2021).

On the other hand, since maternal mental health can impair childcare practices and ECD (Patel et al., 2004), an intervention targeting maternal mental health was effective in increasing time-intensive investment in children (Bhalotra et al., 2020), as well as improving mothers' mental health and children's cognitive development when mental support and ECD treatments were offered in bundle (Singla et al., 2015).

Intergenerational transmission of mental health. We also contribute to a small but growing literature on the intergenerational transmission of mental health. The previous focus has been on the transmission of health from the older generation to the new, with the mechanism being that various genetic and environmental factors can make the newer generation susceptible to various diseases (Ahlburg, 1998). In contrast, the channels that allow the transmission of mental health from parents to children are the connectedness, care, and communication between the two (Ackard et al., 2006). Studies have long used longitudinal survey data to show that parents' mental health is positively correlated with their children's mental health and economic outcomes (Johnston et al., 2013; Eyal & Burns, 2019), can predict poor mental health among daughters (Gonçalves et al., 2016), and increase the take-up of ADHD, anxiety, and depression medication among children (Persson & Rossin-Slater, 2018).

Early-life interventions and adverse life experiences. More broadly, our study also relates to the literature on the importance of early-life interventions on child development and human capital accumulation (Miguel & Kremer, 2004; Bharadwaj et al., 2013; Alan et al., 2021; Carneiro et al., 2021). We also contribute to the literature on the negative consequences of adverse life experiences during childhood, such as due to conflict (Minoiu & Shemyakina, 2012), war (Singhal, 2019), and human-made disasters (Almond et al., 2009).

Table C1: Impact evaluations on mental health

STUDY	Type	Sample	Dosage	Findings
Rahman et al. (2008)	CBT	903 (women), 463 got treated	16 sessions (no data on duration)	↓Depression
Bhalotra et al. (2020)	Follow up of Rahman et al. (2008)	585 (women)	-	↓Depression
Barker et al. (2022)	CBT	7,227 (adults), 1,290 got treated	12 sessions (18 hours)	↓Distress
McKelway et al. (2022)	CBT	1,120 (elderly), 376 got treated	6 sessions (3-4.5 hours)	↓Depression

[4] Andrew et al. (2018), however, did not find a persistent impact of Attanasio et al. (2014) on ECD outcomes two years later.

Patel et al. (2017)	CBT/BA	495 (adults), 247 got treated	6-8 sessions (3-5 hours)	↓Depression
Fuhr et al. (2019)	CBT/BA	280 (adults), 140 got treated	6-14 sessions (3-10.5 hours)	↓Depression
Bhat et al. (2022)	Follow up of Patel et al. (2017) Fuhr et al. (2019)	493 (adults) + 280 (adults)	-	↓Depression, but only Patel et al. (2017)
Maselko et al. (2020)	CBT	570 (women), 284 got treated	18 sessions (no data on duration)	No effect
Tol et al. (2020)	ACT	694 (refugees), 331 got treated	5 sessions (10 hours)	↓Depression ↓Trauma
Bryant et al. (2017)	PM+	421 (women), 209 got treated	5 sessions (7.5 hours)	↓Depression
Haushofer et al. (2020)	PM+	5,756 (adults), 525 got PM+	5 sessions (7.5 hours)	No effect
Acarturk et al. (2022)	PM+	46 (refugees), 24 got treated	5 sessions (7.5 hours)	↓Distress ↓Trauma
de Graaff et al. (2020)	PM+	60 (refugees), 30 got treated	5 sessions (7.5 hours)	↓Depression ↓Trauma ↓Anxiety
Rahman et al. (2019)	PM+	612 (women), 306 got treated	5 sessions (10 hours)	↓Depression ↓Trauma ↓Anxiety
Rahman et al. (2016)	PM+	346 (women), 172 got treated	5 sessions (7.5 hours)	↓Depression ↓Trauma ↓Anxiety
Bolton et al. (2003)	ITP	341 (adults), 163 got treated	16 sessions (24 hours)	↓Depression
Christensen et al. (2004)	Psychoeducation	525 (adults), 165 got treated	5 sessions (no data on duration)	↓Depression
Geisner et al. (2006)	Psychoeducation	177 (adults), 89 got treated	1 session (no data on duration)	↓Depression
Vlassopoulos et al. (2024)	Informational	2,402 (adults), 1,299 got treated	4 sessions (2 hours)	↓Depression ↓Stress
Sadish et al. (2021)	Informational	914 (adults), no data on how many got treated	Once, over-the-phone no data on duration	↓Depression ↓Anxiety

Note: CBT is Cognitive Behavioral Therapy; BA is Behavioral Activation; ACT is Acceptance and Commitment Therapy, a modern variant of CBT; PM+ is Problem Management Plus; ITP is interpersonal psychotherapy.

C.2 Heterogeneity analysis using interactions

To estimate whether treatment effects vary by children's gender, household's exposure to violence during the conflict, mother's experiences of abuse in the camp, mother education, and age of mothers and children, we estimate the following interaction model:

$$Y_{1ijc} = \beta_0 + \beta_1 Treat_{jc} + \beta_2 G_{ijc} + \beta_3 Treat_{jc} \times G_{ijc} + \beta_4 Y_{0ijc} + \mathbf{\Gamma}' \boldsymbol{X}_{ijc} + \theta_c + \epsilon_{ijc} \quad (2)$$

where G_{ijc} is either children's gender (an indicator for male), an indicator for high exposure to violence during the conflict in Myanmar (=1 if the household victimization index is above the median value and 0 otherwise), an indicator for more experience of abuse in the camp (=1 if more and 0 otherwise), and an indicator for mother that attended primary school (=1 if true and 0 otherwise), and an indicator for old (=1 if mother/child's age is above the median value and 0 otherwise).

We did not have a specific hypothesis regarding the potential impact of the intervention on the mental health of mothers with male versus female children under 2 years of age. It is possible that the treatment had a greater influence on the mental health of mothers with sons, as studies suggest that parents tend to be more satisfied and optimistic when they have male children rather than female ones (Raley & Bianchi, 2006). Additionally, the prevalence of son preference may have caused mothers to be more attentive and engaged during counseling sessions, resulting in different outcomes. Similarly, mothers and children from households that experienced greater levels of violent conflict in Myanmar may have been more strongly affected by our intervention than those from households with less exposure to violence, as traumatic memories are likely to be more frequent among those highly exposed.

Table C2 presents the heterogeneity results by children's gender and household exposure to violence in Myanmar. Column 1 reports the pooled effects (same as column 2 in Table 2 in the main paper), 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), development (Panel B2), and anthropometric (Panel B3) 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 children's gender.

We will now investigate how exposure to violence affects the results. Specifically, we will examine estimates for highly exposed individuals in Column 5 and compare them to estimates for the least exposed in Column 6. Column 7 shows the coefficients for the interaction term. When looking at mental health outcomes (Panels A1 and B1), our analysis reveals that the treatment effects are greater for highly exposed individuals compared to those who are least exposed. However, for mothers' mental health, the differences are only slightly significant. This suggests that the improvement in mental health for highly exposed mothers is more significant than the improvement observed for those with low exposure to violent conflict. Interestingly, we also found that the treatment effect on mothers' aspirations varies based on violence exposure, with aspirations of highly exposed mothers deteriorating more after the intervention. However, this difference is only marginally significant ($p < 0.10$). We did not find any evidence of heterogeneity by violence exposure regarding children's development and anthropometric outcomes.

We also conduct additional heterogeneity analysis by experiences of abuse by mothers in the refugee camp and mothers' education level. These results are reported in Table C3. We do not find any heterogeneity in mental health impacts (neither of mothers nor children) by camp-based abuse and education level. In the case of development, we find that mothers that did not encounter any camp abuse, their children experienced a significant improvement in problem-solving skills than children of mothers that encountered at least one camp abuse (column 4, Panel B2). In addition, children of uneducated mothers (i.e., never went to primary school) benefited the most in terms of improvements in communication and personal-social skills (column 7, Panel B2).

Finally, we explore heterogeneous treatment effects by age of mothers and children in Table C4. We only find that older children benefited the most in terms of improvements in communication

skills. Moreover, although marginally significant, we also find that younger children benefited more in terms of trauma reductions and younger mothers benefited more in terms of improvements in their sense of belongingness. For the rest, we do not observe any heterogeneity by age.

Table C2: 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' well-being							
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)
B3. Children's anthropometrics							
Height-for-age z-score	0.515***	0.645***	0.417**	0.015	0.530***	0.541***	-0.064
	(0.139)	(0.192)	(0.193)	(0.256)	(0.195)	(0.195)	(0.281)
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 are derived from the sample with only female children and column 3 are from male children sample. High exposure=1 when households' exposure to violence in Myanmar is higher than the median value and 0 if low. Columns 4 and 7 report the coefficients on the interaction term from equation 2.

Table C3: Heterogeneous treatment effects, by mothers' camp abuse and education

		Victim of at least one camp abuse			Attended primary school		
	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	-0.218***	-0.267***	0.089
	(0.051)	(0.120)	(0.052)	(0.101)	(0.056)	(0.078)	(0.079)
Depression	-0.144***	-0.206**	-0.133**	-0.015	-0.153**	-0.116	-0.008
	(0.054)	(0.104)	(0.053)	(0.084)	(0.061)	(0.072)	(0.069)
A2. Mothers' well-being							
Happiness	0.117**	0.090	0.118**	0.016	0.157**	0.006	0.095
	(0.056)	(0.105)	(0.058)	(0.101)	(0.063)	(0.086)	(0.084)
Aspirations	-0.066	-0.160*	-0.047	-0.072	-0.106	0.038	-0.168*
	(0.062)	(0.096)	(0.068)	(0.105)	(0.066)	(0.100)	(0.097)
Belongingness	0.179***	0.322***	0.153***	0.175	0.192***	0.171*	0.032
	(0.057)	(0.110)	(0.058)	(0.111)	(0.060)	(0.101)	(0.094)
B1. Children's mental health							
Trauma	-0.096*	-0.057	-0.107*	0.130	-0.095	-0.086	0.028
	(0.055)	(0.110)	(0.058)	(0.101)	(0.063)	(0.077)	(0.089)
Depression	-0.122**	-0.210*	-0.103*	-0.074	-0.109	-0.164**	0.043
	(0.059)	(0.119)	(0.061)	(0.111)	(0.069)	(0.081)	(0.094)
B2. Children's development							
Communication skills	0.229***	0.253**	0.230***	-0.025	0.177***	0.412***	-0.233***
	(0.059)	(0.100)	(0.062)	(0.090)	(0.064)	(0.092)	(0.090)
Gross-motor skills	0.179***	0.156	0.186***	-0.087	0.142**	0.276***	-0.138
	(0.058)	(0.101)	(0.061)	(0.101)	(0.062)	(0.103)	(0.097)
Fine-motor skills	-0.021	0.035	-0.023	-0.055	-0.059	0.101	-0.127
	(0.066)	(0.118)	(0.069)	(0.107)	(0.068)	(0.117)	(0.105)
Problem-solving skills	0.177***	0.037	0.206***	-0.249**	0.148**	0.268**	-0.098
	(0.055)	(0.115)	(0.056)	(0.110)	(0.059)	(0.104)	(0.100)
Social skills	0.128*	0.209*	0.120*	-0.022	0.068	0.330***	-0.225**
	(0.067)	(0.121)	(0.071)	(0.120)	(0.069)	(0.116)	(0.103)
B3. Children's anthropometrics							
Height-for-age z-score	0.515***	0.865**	0.461***	0.427	0.571***	0.456*	0.195
	(0.139)	(0.393)	(0.152)	(0.386)	(0.167)	(0.261)	(0.287)
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 2.

Table C4: Heterogeneous treatment effects, by age

		Mothers' age			Children's age		
	Pooled	Old	Young	Difference (β_3)	Old	Young	Difference (β_3)
Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A1. Mothers' mental health							
Trauma	-0.233***	-0.249***	-0.210***	-0.037	-0.229***	-0.225***	0.014
	(0.051)	(0.062)	(0.060)	(0.068)	(0.063)	(0.059)	(0.063)
Depression	-0.144***	-0.184***	-0.093	-0.091	-0.102	-0.187***	-0.022
	(0.054)	(0.065)	(0.062)	(0.066)	(0.069)	(0.065)	(0.061)
A2. Mothers' well-being							
Happiness	0.117**	0.063	0.179***	-0.115	0.125*	0.105	0.001
	(0.056)	(0.067)	(0.067)	(0.071)	(0.064)	(0.068)	(0.073)
Aspirations	-0.066	-0.051	-0.071	0.018	-0.105	-0.033	-0.078
	(0.062)	(0.073)	(0.078)	(0.081)	(0.076)	(0.074)	(0.080)
Belongingness	0.179***	0.094	0.281***	-0.127*	0.142**	0.209***	-0.030
	(0.057)	(0.067)	(0.069)	(0.073)	(0.070)	(0.065)	(0.072)
B1. Children's mental health							
Trauma	-0.096*	-0.086	-0.114	-0.005	-0.040	-0.155**	0.132*
	(0.055)	(0.061)	(0.074)	(0.076)	(0.062)	(0.069)	(0.069)
Depression	-0.122**	-0.106	-0.144**	0.039	-0.137*	-0.100*	-0.001
	(0.059)	(0.069)	(0.070)	(0.071)	(0.077)	(0.060)	(0.071)
B2. Children's development							
Communication skills	0.229***	0.291***	0.153**	0.108	0.371***	0.089	0.221**
	(0.059)	(0.064)	(0.076)	(0.073)	(0.061)	(0.085)	(0.088)
Gross-motor skills	0.179***	0.189***	0.170**	0.022	0.246***	0.118	0.074
	(0.058)	(0.064)	(0.075)	(0.070)	(0.066)	(0.081)	(0.087)
Fine-motor skills	-0.021	0.010	-0.053	0.063	-0.041	-0.004	-0.080
	(0.066)	(0.072)	(0.081)	(0.076)	(0.084)	(0.074)	(0.089)
Problem-solving skills	0.177***	0.190***	0.164**	0.043	0.254***	0.109	0.117
	(0.055)	(0.067)	(0.069)	(0.076)	(0.071)	(0.069)	(0.087)
Social skills	0.128*	0.163**	0.085	0.049	0.244***	0.005	0.125
	(0.067)	(0.077)	(0.076)	(0.074)	(0.076)	(0.087)	(0.098)
B3. Children's anthropometrics							
Height-for-age z-score	0.515***	0.630***	0.379*	0.311	0.213	0.871***	-0.434
	(0.139)	(0.188)	(0.203)	(0.259)	(0.186)	(0.204)	(0.273)
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 2. Old=1 if age is higher than the median (25 years of mothers and 14 months for children) and 0 otherwise.

References

Acarturk, C., Uygun, E., Ilkkursun, Z., Yurtbakan, T., Kurt, G., Adam-Troian, J., Senay, I., Bryant, R., Cuijpers, P., Kiselev, N., et al. (2022). Group problem management plus (pm+) to decrease psychological distress among syrian refugees in turkey: a pilot randomised controlled trial. *BMC Psychiatry*, *22*(1), 1–11.

Ackard, D. M., Neumark-Sztainer, D., Story, M., & Perry, C. (2006). Parent–child connectedness and behavioral and emotional health among adolescents. *American Journal of Preventive Medicine*, *30*(1), 59–66.

Ahlburg, D. (1998). Intergenerational transmission of health. *American Economic Review Papers & Proceedings*, *88*(2), 265–270.

Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *Quarterly Journal of Economics*, *136*(4), 2147–2194.

Almond, D. & Currie, J. (2011). Human capital development before age five. In *Handbook of Labor Economics*, volume 4 (pp. 1315–1486). Elsevier.

Almond, D., Edlund, L., & Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, *124*(4), 1729–1772.

Amadu, S., Attanasio, O. P., Caeyers, B., Cattan, S., Sosa, L. C., Krutikova, S., Leighton, P., Masselus, L., & Yakubu, M. (2019). Improving early childhood development in rural Ghana through scalable low-cost community-run play schemes: Programme impact evaluation report. Technical report, IFS Report.

Andrew, A., Attanasio, O., Fitzsimons, E., Grantham-McGregor, S., Meghir, C., & Rubio-Codina, M. (2018). Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in colombia. *PLoS Medicine*, *15*(4), e1002556.

Attanasio, O., Baker-Henningham, H., Bernal, R., Meghir, C., Pineda, D., & Rubio-Codina, M. (2022). Early stimulation and nutrition: The impacts of a scalable intervention. *Journal of European Economic Association (forthcoming)*.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., & Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *BMJ*, *349*.

Barker, N., Bryan, G. T., Karlan, D., Ofori-Atta, A., & Udry, C. R. (2022). Cognitive Behavioral Therapy Among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress. *Forthcoming at American Economic Review: Insights*.

Baumgartner, J. N., Ali, M., Gallis, J. A., Lillie, M., Owusu, R., Abubakr-Bibilazu, S., Adam, H., Aborigo, R., McEwan, E., Zhou, Y., et al. (2021). Effect of a lay counselor-delivered integrated maternal mental health and early childhood development group-based intervention in northern ghana: a cluster-randomized controlled trial. *Global Mental Health*, *8*.

Bhalotra, S., Baranov, V., Biroli, P., & Maselko, J. (2020). Maternal depression, women's empower-

ment, and parental investment: Evidence from a randomized controlled trial. *American Economic Review*, *110*(3), 824–59.

Bharadwaj, P., Løken, K. V., & Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, *103*(5), 1862–91.

Bhat, B., de Quidt, J., Haushofer, J., Patel, V. H., Rao, G., Schilbach, F., & Vautrey, P.-L. P. (2022). The long-run effects of psychotherapy on depression, beliefs, and economic outcomes. *NBER working paper*.

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.

Bryant, R. A., Schafer, A., Dawson, K. S., Anjuri, D., Mulili, C., Ndogoni, L., Koyiet, P., Sijbrandij, M., Ulate, J., Harper Shehadeh, M., et al. (2017). Effectiveness of a brief behavioural intervention on psychological distress among women with a history of gender-based violence in urban kenya: a randomised clinical trial. *PLoS Medicine*, *14*(8), e1002371.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485.

Campbell, F. A. & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Development*, *65*(2), 684–698.

Carneiro, P., Kraftman, L., Mason, G., Moore, L., Rasul, I., & Scott, M. (2021). The impacts of a multifaceted prenatal intervention on human capital accumulation in early life. *American Economic Review*, *111*(8), 2506–49.

Christensen, H., Griffiths, K. M., & Jorm, A. F. (2004). Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ*, *328*(7434), 265.

Cronin, C. J., Forsstrom, M. P., & Papageorge, N. W. (2020). What good are treatment effects without treatment? mental health and the reluctance to use talk therapy. *NBER working paper no. 27711*.

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? a meta-analytic update of the evidence. *World Psychiatry*, *15*(3), 245–258.

Cuijpers, P., Muñoz, R. F., Clarke, G. N., & Lewinsohn, P. M. (2009). Psychoeducational treatment and prevention of depression: the "coping with depression" course thirty years later. *Clinical Psychology Review*, *29*(5), 449–458.

Currie, J. & Stabile, M. (2007). Mental health in childhood and human capital. In *The problems of disadvantaged youth: An economic perspective* (pp. 115–148). University of Chicago Press.

de Graaff, A. M., Cuijpers, P., McDaid, D., Park, A., Woodward, A., Bryant, R. A., Fuhr, D. C., Kieft, B., Minkenberg, E., & Sijbrandij, M. (2020). Peer-provided problem management plus (pm+) for adult syrian refugees: a pilot randomised controlled trial on effectiveness and cost-effectiveness. *Epidemiology and Psychiatric Sciences*, *29*.

Eyal, K. & Burns, J. (2019). The parent trap: cash transfers and the intergenerational transmission of depressive symptoms in South Africa. *World Development*, *117*, 211–229.

Fuhr, D. C., Weobong, B., Lazarus, A., Vanobberghen, F., Weiss, H. A., Singla, D. R., Tabana, H., Afonso, E., De Sa, A., D'Souza, E., et al. (2019). Delivering the thinking healthy programme

for perinatal depression through peers: an individually randomised controlled trial in india. *The Lancet Psychiatry*, *6*(2), 115–127.

Geisner, I. M., Neighbors, C., & Larimer, M. E. (2006). A randomized clinical trial of a brief, mailed intervention for symptoms of depression. *Journal of Consulting and Clinical Psychology*, *74*(2), 393.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., & Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, *344*(6187), 998–1001.

Gertler, P., Heckman, J. J., Pinto, R., Chang, S. M., Grantham-McGregor, S., Vermeersch, C., Walker, S., & Wright, A. (2021). Effect of the jamaica early childhood stimulation intervention on labor market outcomes at age 31. *NBER working paper*.

Gonçalves, H., Pearson, R., Horta, B., González-Chica, D., Castilho, E., Damiani, M., Lima, R., Gigante, D., Barros, F., Stein, A., et al. (2016). Maternal depression and anxiety predicts the pattern of offspring symptoms during their transition to adulthood. *Psychological Medicine*, *46*(2), 415–424.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., & Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the jamaican study. *The Lancet*, *338*(8758), 1–5.

Haushofer, J., Mudida, R., & Shapiro, J. (2020). The comparative impact of cash transfers and psychotherapy on psychological and economic well-being. *NBER working paper no. 28106*.

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.

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.

Hussam, R. N., Kelley, E. M., Lane, G. V., & Zahra, F. T. (2022). The psychosocial value of employment: Evidence from a refugee camp. *American Economic Review*, *112*(11), 3694–3724.

Johnston, D. W., Schurer, S., & Shields, M. A. (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of Health Economics*, *32*(6), 1077–1089.

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.

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

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

Maselko, J., Sikander, S., Turner, E. L., Bates, L. M., Ahmad, I., Atif, N., Baranov, V., Bhalotra, S., Bibi, A., Bibi, T., et al. (2020). Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: a cluster randomised trial in pakistan. *The Lancet Psychiatry*, *7*(9), 775–787.

McKelway, M., Banerjee, A., Grela, E., Schilbach, F., Sharma, G., Sequeira, M., Vaidyanathan, G., & Duflo, E. (2022). Impacts of cognitive behavioral therapy and cash transfers on depression and impairment of elderly living alone: A randomized trial in india. *NBER working paper*.

Miguel, E. & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217.

Minoiu, C. & Shemyakina, O. (2012). Child health and conflict in Côte d'Ivoire. *American Economic Review*, *102*(3), 294–99.

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.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., et al. (2017). The healthy activity program (hap), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in india: a randomised controlled trial. *The Lancet*, *389*(10065), 176–185.

Persson, P. & Rossin-Slater, M. (2018). Family ruptures, stress, and the mental health of the next generation. *American Economic Review*, *108*(4-5), 1214–52.

Rahman, A., Hamdani, S. U., Awan, N. R., Bryant, R. A., Dawson, K. S., Khan, M. F., Azeemi, M. M.-U.-H., Akhtar, P., Nazir, H., Chiumento, A., et al. (2016). Effect of a multicomponent behavioral intervention in adults impaired by psychological distress in a conflict-affected area of pakistan: a randomized clinical trial. *JAMA*, *316*(24), 2609–2617.

Rahman, A., Khan, M. N., Hamdani, S. U., Chiumento, A., Akhtar, P., Nazir, H., Nisar, A., Masood, A., Din, I. U., Khan, N. A., et al. (2019). Effectiveness of a brief group psychological intervention for women in a post-conflict setting in pakistan: a single-blind, cluster, randomised controlled trial. *The Lancet*, *393*(10182), 1733–1744.

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.

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

Reynolds, C. F., Dias, A., Cohen, A., Morse, J., Anderson, S. J., Cuijpers, P., & Patel, V. (2017). Preventing late-life depression: lessons in intervention development from goa, india. *Innovation in Aging*, *1*(3), igx030.

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

Sadish, D., Adhvaryu, A., & Nyshadham, A. (2021). (mis) information and anxiety: Evidence from a randomized covid-19 information campaign. *Journal of Development Economics*, *152*, 102699.

Singhal, S. (2019). Early life shocks and mental health: The long-term effect of war in Vietnam. *Journal of Development Economics*, *141*, 102244.

Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: lessons from low-and middle-income countries: annual review of clinical psychology. *Annual Review of Clinical Psychology*, *13*, 149.

Singla, D. R., Kumbakumba, E., & Aboud, F. E. (2015). Effects of a parenting intervention to address maternal psychological wellbeing and child development and growth in rural uganda: a community-based, cluster-randomised trial. *The Lancet Global Health*, *3*(8), e458–e469.

Tol, W. A., Leku, M. R., Lakin, D. P., Carswell, K., Augustinavicius, J., Adaku, A., Au, T. M., Brown, F. L., Bryant, R. A., Garcia-Moreno, C., et al. (2020). Guided self-help to reduce psychological distress in south sudanese female refugees in uganda: a cluster randomised trial. *The Lancet Global Health*, *8*(2), e254–e263.

Vlassopoulos, M., Siddique, A., Rahman, T., Pakrashi, D., Islam, A., & Ahmed, F. (2024). Improving women's mental health during a pandemic. *American Economic Journal: Applied Economics*, *16*(2), 422–455.

Walker, S. P., Chang, S. M., Vera-Hernández, M., & Grantham-McGregor, S. (2011). Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics*, *127*(5), 849–857.

Yousafzai, A. K., Obradović, J., Rasheed, M. A., Rizvi, A., Portilla, X. A., Tirado-Strayer, N., Siyal, S., & Memon, U. (2016). Effects of responsive stimulation and nutrition interventions on children's development and growth at age 4 years in a disadvantaged population in pakistan: a longitudinal follow-up of a cluster-randomised factorial effectiveness trial. *The Lancet Global Health*, *4*(8), e548–e558.

Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2014). Effect of integrated responsive stimulation and nutrition interventions in the lady health worker programme in pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *The Lancet*, *384*(9950), 1282–1293.