# Caregiver Mental Health and Early Childhood Development in Conflict-Affected Settings: Experimental Evidence from Colombia

Andrés Moya* María José Torres Juliana Sánchez
Arturo Harker Alicia Lieberman Blasina Niño Vilma Reyes

February 28, 2024

**Abstract**

We report the results from a cluster-randomized control trial of a community and group-based psychosocial program that seeks to promote caregiver mental health as an outcome and as a pathway to protect young children from the effects of conflict and forced displacement. We implemented the program in Colombia with caregivers who have suffered conflict-related violence or have been forcefully displaced. Eight months after the program ended, we found positive intent-to-treat effects of 0.14 standard deviations (sd) on caregiver mental health; 0.27 sd on the quality of child-caregiver interactions; and 0.17 and 0.21 sd on early childhood mental health and development, respectively. Our findings speak to the importance of addressing caregivers' mental health as a *psychological constraint* in early childhood programming in conflict-affected settings and other environments of extreme adversity.

JEL Codes: O15, O12, I15, J13, J24

*Corresponding Author, Facultad de Economía, Universidad de los Andes, a.moya@uniandes.edu.co. Torres, Sánchez, Harker, and Niño: Universidad de los Andes. Lieberman and Reyes: Child Trauma Research Program, University of California, San Francisco. We thank Orazio Attanasio, Raquel Bernal, Jake Bowers, Michael R. Carter, Arnold Cecile, Jorge Cuartas, Fabio Idrobo, Rachid Laajaj, Dana McCoy, Joel Middleton, Fernando Saltiel, Pieter Serneels, Anuraj Snahkar, Hiro Yoshikawa, and participants at seminars at ASSA, Boston College, EAFIT, IFPRI, Harvard, LACEA, NYU, Ridge Impact Evaluation Network, UC Davis, Universidad de los Andes, University of Wisconsin, Madison, and Yale University for valuable comments. We are grateful to Isabella Caro, Daniel Pinzón, and Fayber Acosta, our research assistants, the community facilitators and supervisors, the enumerators and supervisors from Iquartil, and especially to the caregivers and their children who participated in the evaluation. This impact evaluation was approved by the Ethics Committee (IRB) of Universidad de los Andes (Protocol 835 - April 05, 2018) and pre-registered at the AEA RCT Registry (Protocol 2868) on April 09, 2018 and at the Clinical Trials Registry (Protocol NCT03502252) on April 18, 2018 before the program's implementation and baseline data collection. An amendment was filed at the AEA RCT registry on April 20, 2020, to account for the disruption in data collection caused by the COVID-19 pandemic. We acknowledge generous financial support the Saving Brains partnership (Grant Agreement ID SB-POC-1809-19091), United Way Colombia, Fundación Éxito, Fundación FEMSA, Genesis Foundation, Fundación Coca-Cola, Primero lo Primero, and the School of Economics at Universidad de los Andes.

# 1 Introduction

A large body of interdisciplinary work emphasizes the importance of early childhood for brain development, skill formation, and health, which lay the foundations for lifelong socioeconomic trajectories and wellbeing (Currie and Almond, 2011; Black et al., 2017; Knudsen et al., 2006). To reach their full potential, young children need stable environments, enriching experiences, and nurturing relationships with an adult caregiver (Britto et al., 2017).

Unfortunately, millions of young children across the world are at risk of not reaching their developmental potential due to the prevalence of wars, armed conflicts, and forced displacement. By the end of 2020, one in five children under 5 years of age lived in an area affected by armed conflict, while children make up 40 percent of the displaced and refugee populations worldwide. In these settings, the most important factor determining children's resilience and early childhood development is the quality of the relationship with an adult caregiver. However, exposure to traumatic events also affects caregivers' mental health, which can then impair their capacity to provide the nurturing care that their children need (Sánchez Ariza et al., 2023). The combination of early life trauma and a lack of nurturing care can lead to an over-activation of different biological systems, including the body's stress-response system, a phenomenon frequently referred to as *toxic stress*. This leads to potentially permanent changes in brain architecture, neural threat-response, and emotional processing, which hinder healthy early childhood development (Shonkoff et al., 2012; McLaughlin et al., 2019; Nelson and Gabard-Durnam, 2020). Although the consequences of early life adverse experiences are well understood, the needs of young children and their caregivers have been rarely addressed in conflict-affected settings and among forcibly displaced populations, while our understanding of the nexus between caregivers' mental health and early childhood development in these settings is limited.

To contribute towards closing these gaps, we report the results of the experimental evaluation of *Semillas de Apego*, a community and group-based psychosocial program for

caregivers of young children affected by conflict or forced migration. The program is implemented over 15 weekly sessions in groups of 15 to 20 caregivers of young children that are led by two para-professional community facilitators. The program focuses on restoring caregivers' mental health as an outcome and as a pathway to promoting healthy and nurturing relationships, which are essential to moderate children's stress response and prevent dysregulation of different biological systems, thereby protecting children's early development amid adverse or traumatic experiences (Lieberman and Van Horn, 2011). As we discuss in more detail below, this approach differs from that of other successful parenting programs by recognizing that in settings of extreme adversity and trauma, caregivers' mental health is a *psychological constraint* that undermines their capacity to provide nurturing care for their children and leaves young children at risk.

We set our study in Colombia, a country ravaged by decades of ongoing conflict where civilians have been disproportionately affected. By the end of 2023, 8.6 million persons had been internally displaced by conflict in this country, representing 18 percent of it's population and the highest figure worldwide (UNHCR, 2022; Registro Único de Víctimas, 2023).[1] Within Colombia, we conducted the impact evaluation in Tumaco, a municipality that has been historically affected by conflict and forced displacement, and where the majority of its population has suffered, either directly or indirectly, from conflict-related violence and displacement, as we describe in Section 3. Baseline data collected from the caregivers invited to participate in the impact evaluation highlights the pervasiveness of conflict in this context: 82 percent of them had been directly victimized at least once in the previous 10 years, 54 percent had been forcibly displaced at least once over the same period, and over 40 percent had at-risk symptoms of mental health problems.

For the impact evaluation, we targeted mothers or other primary caregivers of children aged 2 to 5 served by public childcare centers serving vulnerable and under-served families. We designed the experimental evaluation to overcome the challenges brought forth by the ongoing conflict and the resource-constrained setting. Some of the elements in our

[1] Colombia is also the largest host of Venezuelan displaced migrants with 2.9 million Venezuelans.

design include (1) the random assignment of the treatment at the childcare center level to overcome the 'invisible frontiers' set up by armed groups that prevent hosting sessions with caregivers from different neighborhoods; (2) the implementation of the model with a small number of trained community facilitators and across four sequential cohorts to reach the sample size required by our power calculations; (3) the collection of all data at the childcare centers, which provided a safe space for participants and enumerators. In making these design decisions, we prioritized ensuring the safety of participants, enumerators, and our implementation and research teams, while considering the possible trade-offs in terms of statistical power and data collection. We discuss our experimental design and these challenges in detail in Section 4.

Over four sequential cohorts, 1,372 caregivers participated in the impact evaluation, with 714 assigned to the treatment group and 662 to control. Caregivers in the control group received regular services for their children and themselves from the childcare centers, while caregivers in the treatment group were invited to participate in the program and continued receiving the regular services provided by the childcare centers. On average, caregivers assigned to the treatment arm participated in 11 out of 15 total sessions, conditional on attending the first session.

We collected data at baseline and 1 and 8-month follow-ups on the five core dimensions of the program following the program's theory of change described in Section 2: caregiver mental health, child-caregiver interactions, quality of the child-caregiver relationships, and children's mental health and early childhood development. For this purpose, we administered a battery of scales and assessments used in previous studies in Colombia that have good psychometric properties in our sample. At baseline, we also administered a standard socioeconomic survey and a module on conflict-related experiences. In Section 5, we describe these scales and their psychometric properties, and the data-collection protocols we implemented to overcome the challenges brought forth by the context.

Our main empirical strategy exploits the random assignment to the treatment at the

childcare center level.[2] For each dimension of interest, we first estimate the program's effects on a summary index following Kling et al. (2007) to reduce the multidimensionality of the data and the underlying noise in each scale. Then, we unpack the results by analyzing each scale and underlying dimension individually. We estimate a lagged dependent variable model where we regress the outcome of interest at the 1 or 8-month follow-up on the random assignment to the program controlling for the outcome at baseline, baseline individual and household characteristics, and a cohort fixed-effect. We estimate our standard errors at the childcare center level following Pustejovsky and Tipton (2018a) to account for the small number of clusters and adjust our estimates for multiple hypothesis testing using Anderson's q-value (Anderson, 2008). Unfortunately, the COVID-19 pandemic and the national lockdown implemented by the Colombian government prevented us from going back to the field to collect the 8-month follow-ups for the last two cohorts of the program. For this reason, the analysis of the impact of the program at the 8-month follow-up uses data from Cohorts 1 and 2 only, as specified in the amendment to our pre-analysis plan in April 2020.

We find that the program led to improvements in caregivers' mental health, the quality and style of child-caregiver interactions, and children's mental health and early development. At the 1-month follow-up, we find positive and statistically significant intent-to-treat (ITT) effects on the style of child-caregiver interactions (0.21 standard deviations (sd)), and on children's mental health and early childhood development (0.14 and 0.20 sd, respectively). At the first follow-up, we also find smaller and not statistically significant effects on caregivers' mental health and the quality of the child-caregiver relationship (0.04 and 0.03 sd, respectively). Yet, at the 8-month follow-up, we find sizeable and statistically significant improvements in the five core dimensions. Specifically, we find ITT effects of 0.14 standard deviations on the caregiver health index; 0.27 and 0.13 sd on the quality and style of child-caregiver interactions, respectively; and 0.17 and 0.21 sd on

[2]We also exploit the model's staggered implementation and the random assignment of participants in the treatment group across four sequential cohorts.

early childhood mental health development. These results suggest that the effects of the program persist and become larger over time, which is important given existing evidence of the fade-out of similar parenting interventions (Bailey et al. 2020).

When we unpack the composite indices, we find positive and sizeable effects across the majority of constructs and sub-dimensions, both at the 1 and 8-month follow-ups. These effects are more salient when assessing the impact on at-risk symptoms, meaning that the program was most effective at the extensive margin. We also find suggestive evidence of heterogeneity of impacts according to two dimensions of baseline vulnerability. On the one hand, we find that the program has stronger effects on caregivers with above-average symptoms of mental health problems at baseline. On the other hand, we find that the program has negligible effects on those who were food-insecure and had baseline levels of wealth below the sample median. Together, these results suggest that the model works best for whom it was designed for (caregivers experiencing mental health problems), but that socioeconomic vulnerabilities should be addressed beforehand or simultaneously to allow caregivers the bandwidth to address their socioemotional needs and capacities, similar to Maslow's hierarchy of needs (Maslow, 1943).

We contribute to different strands of the literature. First, we build upon a large body of work in economics and in other disciplines that analyzes the impacts of parenting programs in low and middle-income countries (LMICs). Implemented through home visits or group meetings, these programs promote positive and responsive parenting by providing information, demonstrations, role-playing, and feedback on caregiver-child interactions (Emmers et al., 2022). Conceptually, this approach addresses caregivers' *cognitive constraints* by allowing caregivers to understand why it is important to promote nurturing care and stimulating environments for their children and how they can provide such care. Experimental evidence from China (Zhou et al., 2022; Luo et al., 2019), Colombia (Attanasio et al., 2022, 2020b) Jamaica (Grantham-McGregor et al., 1991; Gertler et al., 2014), India (Grantham-McGregor et al., 2020; Andrew et al., 2020), and Peru (Araujo et al., 2021) among others, show positive impacts on caregivers' knowledge and behaviors

and early childhood development, with pooled effect sizes between 0.13 to 0.56 standard deviations (Jeong et al., 2021).

Despite this track record, standard parenting programs do not explicitly address caregivers' capacities beyond their cognitive constraints and, in particular, do not consider caregivers' mental health or emotional resources. For instance, a recent review of over 478 early childhood development evaluations, finds that only 12 percent of these studies measure mental health outcomes Evans et al. (2021). When maternal mental health is measured, it is often considered as a dimension of heterogeneity in impact, which is indicative of how standard programming has ignored caregivers' mental health as an outcome and as a key constraint for early childhood development. By contrast, *Semillas de Apego* pioneers by recognizing that conflict, forced displacement, and extreme adversities have a negative toll on caregivers' mental health, which create binding *psychological constraints* that hinder their capacities to provide caring and nurturing environments for their children and that is detrimental for their children's early development. Our results therefore demonstrate the urgency and potential of addressing caregivers' mental health as a way to protect young children affected by conflict and forced displacement. To the best of our knowledge, there is only evidence from one similar approach for Rohingya refugees in Bangladesh (Siddique et al., 2022).

We also contribute to an upcoming body of work in economics that has studied mental health as a core dimension of well-being and as a determinant for human capital formation and socioeconomic trajectories. Work in this area has analyzed the effects of mental health or psychological stress on income, productivity, behavior, and poverty (Haushofer and Fehr, 2014; Moya, 2018; Ridley et al., 2020). Closely related to the theory of change of the program we study, prior work in our discipline has also analyzed the effects of early-life exposure to maternal stress on early childhood physical and mental health and on later human capital accumulation (Camacho, 2008; Fontes et al., 2023; Persson and Rossin-Slater, 2018; Sánchez Ariza et al., 2023). Further, recent work has analyzed the economic and health impacts of mental health interventions demonstrating that tackling

*psychological constraints* can create pathways for human capital accumulation and poverty reduction (Angelucci and Bennett, 2024; Baranov et al., 2020; Bossuroy et al., 2022).

Finally, our results also speak to a large body of work in economics on the socioeconomic consequences of armed conflict, forced displacement, and refugee flows.[3] Our interest as a discipline in these areas is motivated by the surge in conflicts and forced displacement worldwide in the last decade, where 114 million people have been forcibly displaced and over one billion people reside in conflict-affected settings, and by the recognition that these phenomena are a major development challenge. Work in this area has analyzed the effects of providing humanitarian aid, cash transfers, access to education, job training, and regularization programs among others to conflict-affected populations and forcibly displaced persons. Yet, the consequences for the youngest children, aged 0 to 5, who represent over 30 percent of conflict-affected and forcibly displaced populations, have been largely ignored in this area of work. Furthermore, there is a gap in funding and rigorous evidence on programs in settings of ongoing conflict and humanitarian crises (Murphy et al., 2018)[4]. We first contribute to this body of work by demonstrating how the psychological consequences of conflict and forced displacement are an additional mechanism that explains the intergenerational transmission of poverty. Further, we contribute to this body of work by highlighting that it is possible to implement and evaluate psychosocial programs in fragile and conflict-affected settings that help break this cycle of trauma and poverty.

# 2 Conceptual Framework

Semillas de Apego is a psychosocial community-based group program for caregivers of children ages 0-5 who have faced conflict or forced migration. The main objective of

[3] See Blattman and Miguel (2010) seminal work on the causes and consequences of civil conflict (war).

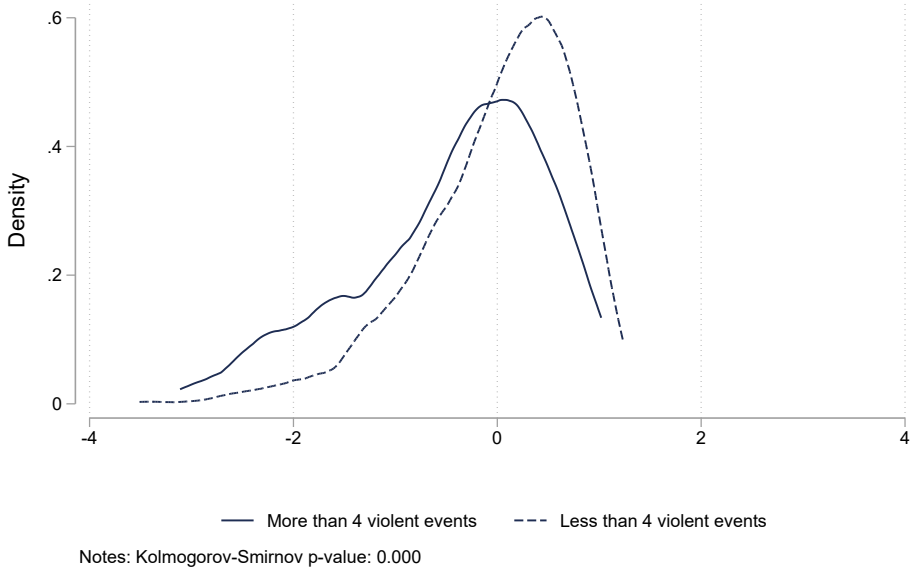
[4] For example, in 2018, educational programs received less than 2 percent of all humanitarian funding and only a small fraction of which is assigned to early childhood development initiatives despite the well-known importance of early childhood development and the overrepresentation of young children among conflict-affected populations (Murphy et al., 2018)

the program is to strengthen the mental health of caregivers and improve the child-parent relationship as a pathway to restore children's mental health and promote their early childhood development. We designed *Semillas de Apego* building from the evidence that adverse circumstances disrupt adults' mental health. Living in a context of ongoing conflict, caregivers must deal with their stress, grief, and emotional world and this hinders their emotional capacity to connect emotionally with their children. This, in combination with the adversities, translates into *toxic stress* and affects young children's mental health and development.

To illustrate this feature, and to provide data that supports the specific approach of the program, in Figures 1 and 2, we use baseline data to describe the patterns of association between the degree of exposure to conflict-related violence and caregiver mental health, child mental health, and early childhood development and the mediating role of caregiver mental health in explaining the effects of conflict on young children.

In Figure 1 we plot the distributions of the caregiver mental health index according to the severity of pre-baseline conflict-related violence, which we stratify arbitrarily according to the sample median. The figure highlights a dose-response relationship between conflict-related violence and mental health; this is, that exposure to more severe conflict is associated with worse mental health, a finding that has been documented in previous studies with IDPs in Colombia (Moya, 2018) and among survivors of mass violence (Mollica et al., 1998).

Figure 1: Conflict and Caregiver Mental Health




*Notes*: The figure plots the distribution of caregivers' mental health according to the severity of the exposure to violent events (above or below the sample median).

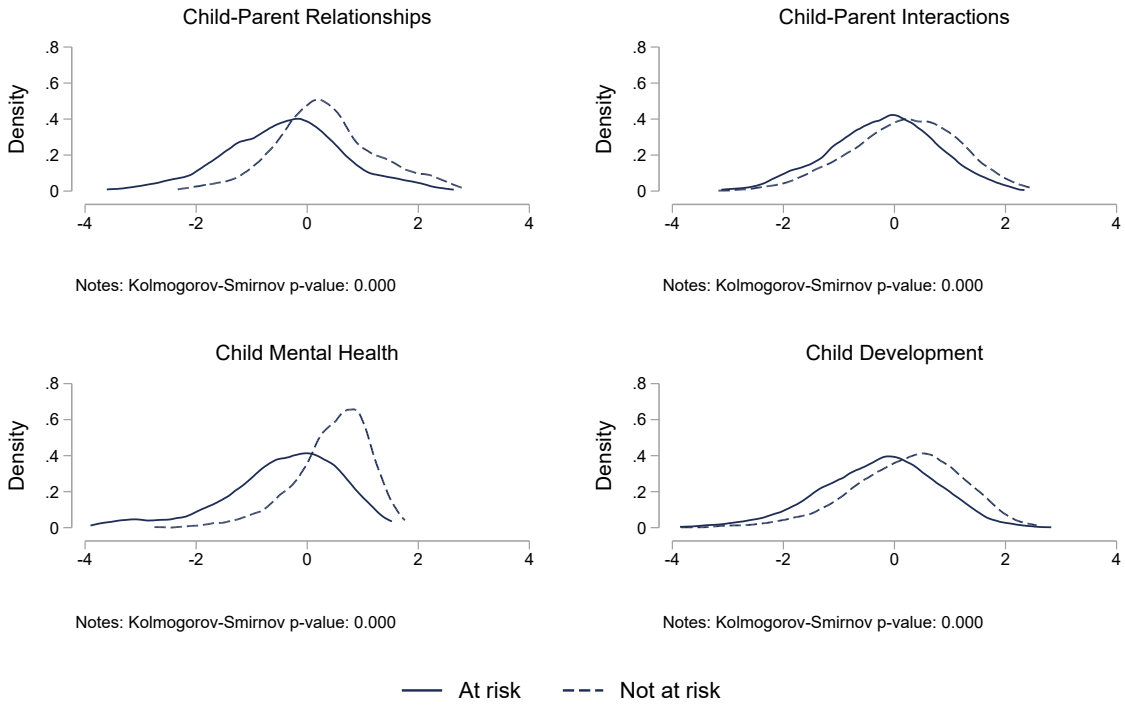
The consequences of these adverse circumstances go beyond caregivers' mental health and hamper their capacity to be emotionally available and to provide nurturing developmentally appropriate, and sensitive caregiving. The lack of a nurturing relationship coupled with the systematic, persistent, and traumatic exposition to stressful circumstances is associated with an over-activation of the biological stress response system. The over-activation of the stress response system during the first five years of life has severe consequences in brain architecture, neural threat response, and emotional processing, which hinder the child's development.

Figure 2 illustrates this. In the figure, we plot the distributions of the four indices that capture the quality and style in the child-caregiver relationship and young children's mental health and development. In these figures, we stratify the data according to whether the caregiver had baseline symptoms of mental health problems.[5] The figures illustrate

[5]As mentioned in the previous section, we consider at-risk symptoms for index scores that are at least one standard deviation below the baseline sample mean.

how caregivers' mental health problems are associated with worse quality and style of child-caregiver relationships and with worse children's mental health and development. Together, this descriptive analysis is indicative of the toll of conflict on caregivers' mental health and of the intergenerational transmission of mental health and poverty in conflict-affected settings and provides further support for the focus of the *Semillas de Apego* program on restoring caregivers' mental health as a way to protect early childhood development. See Sánchez-Ariza et al. (2023) for a more structured regression analysis using this same data, including evidence on the mediating role of caregivers' mental health.

Figure 2: Caregiver Mental Health and Early Childhood Development




*Notes*: The figure plots the distribution of the quality and style of the child-caregiver relationship and of children's mental health and early childhood development according to the caregivers' symptoms of mental health problems (above or below at-risk levels).

Building upon this evidence, *Semillas de Apego* focuses on restoring caregivers' mental health as an outcome and a pathway to promoting healthy and nurturing relationships. These are essential to moderate children's stress response and prevent dysregulation of

different biological systems, thereby protecting children's early development amid adverse or traumatic experiences (Lieberman and Van Horn, 2011).

To promote the caregivers' mental health and improve the child-parent relationship, *Semillas de Apego* builds on three cornerstones that make it scalable and different from other existing programs. First, the program has a psychosocial approach delivered through 15 weekly sessions, focusing on three central themes sequentially. During the first sessions, the program helps caregivers reflect on their life experiences, identify their strengths, and improve their emotional regulation. This promotes their mental health and prepares them to address early childhood development and bonding with their children. The following sessions focus on understanding children's internal processes –what they need during early childhood and how they are affected by adversities— and promoting reflection among participants to provide protection and security for their children, and strengthen the repertoire of parenting strategies, including assertive and loving caregiving. The last sessions address social ecologies in which children develop. Caregivers learn to strengthen their support network and parenting teams. Overall, the program's psychosocial approach and three themes help caregivers provide nurturing and developmentally appropriate caregiving and foster early childhood development. Table 1 summarizes the objectives by dimension, while the Appendix provides a detailed discussion on the curriculum, conceptual framework, activities, and objectives of each session.

Table 1: Semillas de Apego: Main Dimensions and Objectives

| Maternal Mental Health | Early Childhood Development | Child-Parent Relationships | Parenting Teams |
|---|---|---|---|
| Understand the emotional toll of conflict and forced displacement. | Understand the processes of early childhood development and what children need. | Understand how nurturing child-parent relationships protect children from adversities. | Promote parenting teams and restore trust on existing ones. |
| Promote non-judgmental introspection. | Understand how children are affected by adversities. | Understand how adversities affect child-parent relationships. | Engage parenting teams into every-day child-rearing activities. |
| Promote sensory awareness and self-regulation. | Enhance the repertoire of developmentally appropriate parenting strategies. | Restore the capacity to provide sensitive and nurturing care. | Empower caregivers and create distance from gender-specific roles. |
| Recognize strengths and capacities for resilience. | | | |

Second, *Semillas de Apego* is delivered through a community delivery model, wherein the sessions are conducted by a dyad of community facilitators, who may not have any prior experience or training in psychosocial programs. This approach helps to build local capacities to overcome the mental health services deficits that exist in conflict-affected settings and contributes to community buy-in and acceptability. However, it is essential to provide context-appropriate and experiential training to ensure proper implementation. During the one-month training, the facilitators take part in the program as participants and go in-depth into the theoretical framework. This training model allows the theoretical framework to follow the facilitators' personal experiences, ensuring that they go through the personal process and transformation that enables them to better convey the program even though they face the same challenges of living and raising children in conflict-affected settings. At the end of the training, the facilitators implement a small pilot under the guidance of the program's technical supervisors. Throughout the implementation, the facilitators also participate in weekly reflective supervision sessions to focus on their well-being and address their emotional needs, helping to prevent vicarious trauma and burnout. This community approach makes Semillas de Apego feasible in contexts of scarce mental health services and acceptable to the community.

Third, the *Semillas de Apego* program has a group-based component that aims to restore the social connections that armed conflict and trauma often erode. The program is designed for groups of approximately 15 mothers, fathers, and other primary caregivers. The group setting helps participants understand that their responses to abnormal and traumatic events, such as conflict, are normal and shared by others who have experienced similar events. Moreover, the group-based approach encourages the development of new relationship models among group members and the wider community. The group approach helps to rebuild trust, repair relationships, and restore social connections that are often lost due to conflict, migration, and displacement.

# 3 Setting

Colombia has experienced a protracted armed conflict characterized by high levels of civilian victimization and forced displacement. By April 2023, the Colombian State has legally recognized 9.5 million victims of the conflict, with 8.4 million of them being internally displaced persons (IDPs) (Unidad para las Víctimas, 2023). This is the largest number worldwide, representing 19 percent of all IDPs in the world and 16 percent of the Colombian population (United Nations High Commissioner for Refugees, 2021). In addition to those legally recognized as victims, the conflict has had a significant impact throughout the country, especially in regions suitable for coca-leaf crops and drug production and trade, where illegal armed groups have clashed to establish territorial control. See Figure A1 in the Appendix, which illustrates the intensity of forced displacement across Colombian municipalities, as a proxy for the severity of armed conflict, and locates Tumaco, our study site, geographically.

Tumaco is a port city and municipality predominantly of Afro-Colombian population that has been and continues to be heavily affected by the conflict. Its location on the Pacific coast, in a region rich in natural resources, and with a dense network of rivers, provides suitable conditions for coca-leaf crops, drug trade, and illegal mining. Therefore, armed groups have sought control of rural areas where the coca-leaf crops, processing laboratories, and mines are located, and of the port city where cocaine is shipped overseas. To this end, these groups have used violence towards civilians as a deliberate strategy for territorial control, consistent with conflict dynamics throughout the country (Grupo de Memoria Histórica, 2014). The struggle for territorial control intensified after the FARC (Fuerzas Armadas Revolucionarias de Colombia for its Spanish acronym) guerrilla demobilized in 2016 and other armed groups clashed for control of the areas left void, resulting in increased and ongoing civilian victimization and displacement.

Tumaco's conflict and socioeconomic vulnerability are illustrated in Appendix Table A1. In 2018, when we launched the randomized control trial, the homicide rate in Tumaco

was 101 per 100,000 inhabitants, more than three times the national rate, while the intensity of forced displacement was more than 3 times higher than the national average. In addition, 45 percent of the population is below the national multidimensional poverty line, which is more than twice the national average, and only 8 percent of the economically active population has formal employment.

# 4 Experimental Design

In this section, we describe the implementation strategy and experimental design of our study, along with the challenges we faced due to the ongoing conflict. In particular, we had to overcome the existence of 'invisible borders', which are established by armed groups and limit the movement of individuals between and within neighborhoods. As we discuss below, these borders created some tradeoffs for the design and implementation of our study and for ensuring the safety of all participants, facilitators, enumerators, and our research team. We conclude this section by analyzing the experiment's dosage and internal validity.

## 4.1 Implementation Strategy

We implemented the program in partnership with 18 public childcare centers in the urban area of Tumaco.[6] Public childcare centers are the foundation of the 'From Zero to Forever' (*De Cero a 5iempre*) national policy established in 2011 to improve early childhood services for vulnerable families.[7] The centers provide integrated early childhood services, including care, education, nutrition, and health for children aged 2 to 5 following national guidelines, as well as parenting workshops to promote nutritional habits, stimulation, and

[6]These centers were operated by Genesis Foundation, our partner in the field. We excluded one center operated by this partner because it did not have a private space for the group sessions. Six other childcare centers operated by the Catholic Church were not included in the trial.

[7]The *Zero a 5iempre* strategy had positive effects on language and nutrition (Bernal and Ramírez, 2019), but mixed effects when compared to the previous less structured home-based community model (Bernal et al., 2019)

health. The centers in the trial served 1,600 children per year, corresponding to 80 percent of children enrolled in public childcare centers in the urban area of Tumaco. See Appendix Figure A2 for the location of the centers in the evaluation.

We chose to implement the program through the childcare centers because they are safe havens and are protected by armed groups even in the most violent neighborhoods. This allowed us to overcome the 'invisible borders' described above and provided a pathway to reaching children and caregivers in need. In particular, by partnering with the centers, we gained access to the different neighborhoods in Tumaco and were able to design the sampling frame for the experiment using the centers' administrative data. This strategy also allowed us to build upon existing relationships between the centers and the communities, which was essential to build rapport with the caregivers in the centers and promote participation in the trial in a setting where the protracted conflict has eroded social capital and trust. In addition, it allowed us to host the program's sessions at the centers or in the nearby vicinity, which facilitated participation in the program and contributed to ensuring the safety of participants and their children, facilitators, and our research team.[8] However, by partnering with childcare centers we excluded caregivers of younger children (aged 0 to 2 years) who were too young to be served by the centers and who may have benefited more from the promotion of secure attachments due to their developmental stage, which may have affected the power in our experiment.

## 4.2 Experimental Design

1,376 caregivers participated in the study; 714 in the treatment arm and 662 in the control arm. All caregivers in the experiment had access to the child and family services provided by the childcare centers, while caregivers in the treatment arm were also invited to participate in the program.

Enrollment into the experiment was open to mothers or other primary caregivers of

[8]To facilitate participation, we also organized the group sessions around the schedules of caregivers: right after they dropped their children at the centers in the morning or before they picked them up in the afternoon.

children aged 2 to 5 enrolled at the 18 childcare centers regardless of their prior exposure to conflict or mental health needs.[9] We randomized access to the program following a two-stage cluster and staggered design. First, we randomized access to the program at the childcare center level, stratifying by center size, children's age distribution, and the socioeconomic profile of the families served. Second, we used administrative data to randomly assign the caregivers in the treated and control centers to 1 out of 4 implementation cohorts stratifying by child age and sex.[10] In each cohort, caregivers who had been randomly assigned to that cohort were invited to participate in a baseline survey and an assessment of one of their children aged 2 to 5. When a caregiver had more than one child in the age range, we randomly selected one child to participate in the assessments. In treated centers, we then invited caregivers who had completed the survey to join the program's group sessions. We summarize the features of the experimental design in the consort diagram in Appendix Figure A3.

When deciding upon the experimental design, we weighed the challenges brought forth by the ongoing conflict on against the implications of randomizing at the cluster level with a small number of clusters. We ruled out individual-level randomization across the 18 childcare centers because the 'invisible borders' would have put participants in danger had we hosted group sessions with participants from different centers or neighborhoods. We also ruled out individual-level randomization within each center because some centers did not serve enough children and caregivers to create separate treatment and control groups, and because our implementation partner argued against this option for fairness with the families in the same centers and to avoid any pushback from the communities. We therefore chose to randomize access to the treatment at the child-care center level ($N_c = 18$) and considered the implications for statistical power and inference. For the latter, we defined a minimum detectable effect (MDE) of 0.26 sd based on prior early childhood

[9]Because of the prevalence of armed conflict in the study site, we expected that most participants would have been affected by conflict-related violence or forced displacement and would exhibit above-average mental health problems.

[10]Before cohorts 2-4, we updated the randomization lists dropping the caregivers who were no longer affiliated at the centers and adding new caregivers to the end of the lists.

development interventions in Colombia and the region (see for example Attanasio et al. (2020a)) and estimated a sample of 1,280 participants nested in 18 childcare centers to achieve 80 percent power with a significance level of 0.05.[11] In section 7, we discuss how we estimate the clustered standard errors with a small number of clusters.

In turn, the decision to follow a staggered design across four cohorts was driven by our power calculations and the novelty of a structured psychosocial program in the study site. To achieve the desired sample size determined by our power calculations, we needed two implementation cohorts and 12 community facilitators, considering the number of caregivers and children served by each center. However, because the program was being implemented for the first time in the study site and there is stigma around mental health, our implementation partner suggested first assessing whether caregivers were willing to participate in a 15-week psychosocial program before committing to hiring and training a larger group of facilitators. Therefore, we trained six community facilitators to deliver the program to an average of 160 caregivers per cohort. This means that we needed four cohorts to reach our target sample size of 640 caregivers in the treatment group.[12] Implementing the program across cohorts also had two advantages: First, we expected to observe improvements in the quality of the program over time and in the relationships with caregivers and their families throughout the program's implementation. Second, it allowed us to exploit the variation in the timing of treatment within caregivers assigned to the treatment group as an alternative empirical strategy that does not depend on cluster-level randomization (see section 7 for more details on this strategy).

[11] This is the estimate from our power calculations pre-registered before the baseline data were collected that do not account for controlling for prognostic baseline variables and assumes an intra-cluster correlation (ICC) of 0.04. In practice, ICCs are lower, ranging between 0.02 and 0.03 using baseline data and admin data from the childcare centers, which yields an MDE of 0.23 sd. See Appendix Table A5 for ICCs. Further, we will control for baseline prognostic variables to increase power.

[12] As defined in our pre-registry, we organized 9 intervention groups in the first 2 cohorts to build capacities over time and then expanded to 12 groups in the last 2 cohorts. Based on an average of 16 participants per group, we reached an average of 144 participants in the first two cohorts and 192 in each of the second two cohorts.

## 4.3 Dosage

Within the treatment arm, 77 percent of the caregivers invited to participate in the program attended at least one session, while the average attendance was 8.3 out of 15 sessions. See Appendix Table A2 for unconditional and conditional participation over the four cohorts and by cohort.[13] Conditional on attending the first session, dosage increased to 10.8 sessions on average and 64 percent of caregivers who attended the first session went on to participate in 12 or more sessions (see Appendix Figure A4).[14] Taken together, these patterns highlight high levels of participation and acceptability towards the program across the four cohorts and signal to the improvements in the program's quality and in developing relationships of trust with the communities and potential participants. See Harker Roa et al. (2023) for a process evaluation of the program and the factors that contributed to consolidating the program's acceptability.[15]

## 4.4 Internal Validity

We now analyze the validity of our experiment, including the random assignment of the program at the childcare center level, between caregivers in the treatment and control groups, and between caregivers across different cohorts.

First, we use center-level administrative data to show that childcare centers in the treatment and control arms were similar across a set of characteristics including children's age and physical development and the demographic and socioeconomic profile of the families served by the centers (see Appendix Table A4). We note a difference in the number of children enrolled at the centers (a 15 percent difference relative to the average size of

[13] We find that a few baseline characteristics are associated with the likelihood of participating in the program in centers assigned to the treatment arm, but we cannot reject the null hypothesis that baseline characteristics explained participation. See Appendix Table **??**.

[14] As expected, dosage increased after the first implementation Cohort. In Cohort 1, almost 30 percent of caregivers dropped out of the program after the first three sessions and only 28 percent of them attended 12 or more sessions. By contrast, in Cohorts 2 to 4, 73 percent of caregivers who attended the first session went on to participate in 12 or more sessions. See Appendix Figure A5)

[15] Acceptability is one of the outcomes of the implementation science framework that measures how well programs are perceived and received by the target population.

control centers), which is explained by the random assignment of the largest childcare center to the treatment group. However, all centers were run by the same operator, received the same resources from the Colombian government per child served, and provided services of similar quality following national guidelines.

Second, participants in treatment and control groups were similar at baseline across a range of household, caregiver, children, and conflict-related characteristics (see Table 2 in the Appendix). Specifically, we achieved balance at baseline for 19 out of the 24 characteristics we tested for and found that imbalance was explained by caregivers in the treatment arm reporting a higher degree of exposure to conflict-related violence and worse outcomes in the caregiver mental health index at baseline (a 0.11 sd difference). Nevertheless, we fail to reject the null hypothesis that baseline characteristics jointly predict assignment into the treatment (F-test 2.222) as reported in Appendix Table A3. We will isolate baseline differences by controlling for them in the econometric analysis.

Third, we also achieved balance in the random assignment across the four cohorts of the experiment. We find that at baseline, participants in the treatment centers who were assigned to the different cohorts were similar across the same dimensions described above. In Appendix Table A6, we show that baseline characteristics do not jointly predict assignment to a specific cohort (joint-significance F-tests range between 2.35 to 3.62).

Fourth, we rule out contamination in the treatment status, and, given the staggered design, that caregivers in control centers switched to treatment centers in the later cohorts to access the program. Using administrative data for the centers, we observe that less than one percent of the caregivers who were reported in the center-level administrative data in 2018, before the implementation of the first cohort, switched centers in the following cohorts. In the few cases when this happened, we did not invite them to participate in the program or assessments. Furthermore, we find that enrollment rates in Cohorts 2-4 were similar between treatment and control centers, which rules out strategic enrollment; this is, that caregivers who were not served initially by the centers enrolled their children at treatment centers to access the program.

Finally, our experimental design and our data were not designed to test for spillovers from the treatment to the control group or between participants assigned to different cohorts. Because the 'invisible borders' limit the interactions between caregivers of different neighborhoods and childcare centers, the scope of information flows between treatment and control clusters should have been limited, even when some clusters in each experimental arm are in neighborhoods close to each other (see Appendix Figure A2). It is possible, nevertheless, that spillovers emerged across cohorts within treated centers because of information flows between caregivers assigned to different cohorts or from children's peer effects; this is, from children who were assigned to later cohorts and benefited from the interactions with peers whose caregivers participated in earlier cohorts and who could demonstrate improved behavior and regulatory skills. We cannot isolate these potential spillovers in the main empirical strategy, but our estimates of the within-treatment variation strategy should provide a benchmark for such spillovers since they would bias downwards the true impact of the program.

# 5 Data

For each cohort, we collected data at baseline 3 to 4 weeks before the program sessions started and at two follow-ups, 1 and 8 months post-intervention, to assess the immediate impacts of the program and whether they persist over time. Similar to the implementation strategy we discussed above, we designed the field protocols to ensure the safety of caregivers and their children, enumerators, and our research team. Below we describe the data collection protocols and timeline, the data collected, and their validity, including the psychometric properties of the scales and assessments we administered and attrition rates across data-collection waves.

## 5.1 Data-Collection Protocols

The 'invisible borders' also created significant risks for the data collection; not only to caregivers and their children but also to enumerators who could not safely walk around the neighborhoods and visit participants' homes to collect the data. For this reason, enumerators collected all of the data at childcare centers and we developed a security protocol in which they were in daily contact with the neighborhood's community leader and childcare center staff. These individuals assessed the neighborhood's security conditions, sometimes even by contacting members of armed groups, and then guided enumerators and our team through the neighborhoods on their way to the centers when conditions were safe. In addition, we adjusted the timeline for the data collection, switching from a 12-month follow-up as registered in our pre-analysis plan to an 8-month follow-up, to limit the number of caregivers and children we had to track outside of the centers during the second post-intervention assessment.[16]

Collecting data in the centers also provided logistical benefits: we collected the data after caregivers dropped off their children at the centers in the mornings or before they picked them up in the afternoon and we assessed children throughout the school day. However, this strategy limited the scope of data we collected, preventing us from collecting 'objective' observation-based assessments of the quality of the home environment and the quality of the child-caregiver interactions. Our measures in these domains rely on caregiver self-reports using contextually and psychometrically valid scales, as we discuss next.

## 5.2 Outcomes and Measurement

We collected data at baseline and each post-intervention follow-up on the five dimensions of interest according to the program's theory of change: caregiver mental health, style of child-caregiver interactions, quality of the child-caregiver relationship, child mental

[16]Using the centers' administrative data and the baseline for Cohort 1, we estimated that 60 percent of the children would have transitioned out of the centers by the 12-month follow-up. By contrast, at an 8-month follow-up, 85 percent of the children would still be enrolled and in attendance, leaving a more manageable task of contacting the remaining 15 percent outside of the childcare centers.

health, and early childhood development. For this purpose, we administered a battery of psychometric scales and assessments that had been validated or at least translated and used previously in Colombia. Except for the observational assessments of early childhood development, we collected all of the data using scales that rely on caregiver self-reports and could therefore be affected by demand effects. Nevertheless, psychometric scales inquire on multiple items or questions that point to an underlying construct, which helps in controlling the degree of demand effects. For example, the mental health scales we detail below inquire about symptoms associated with mental health problems and trauma in clinical studies, such as headaches, back pains, uneasiness, and difficulties sleeping among others, rather than asking directly about trauma. Furthermore, the structure of the scales allows for analyzing the psychometric properties and validity of the data, as we do in the following subsection. In addition, enumerators were affiliated to and identified with an independent firm that was not associated with the program to minimize demand effects. In Appendix Table A7, we describe each scale and assessment below and provide more information on each instrument, including the main constructs and sub-constructs they capture.

*Caregiver Mental Health*: We administered three scales to assess the mental health of caregivers and the incidence of different psychopathologies. (1) A mental health questionnaire based on the Symptom Checklist 90-R (SCL-90R), which measures symptoms of anxiety, phobic anxiety[17], depression, hostility,[18] and sensitivity[19]; (2) The Locus of Control questionnaire that measures the degree to which individuals believe they control their lives and trajectories; and (3) The Positive Psychological Attitudes (IPPA), which measures life satisfaction and perceived self-efficacy to cope with stressful circumstances.

*Child-Caregiver Relationship*: We administered three scales to measure the quality of the child-caregiver relationship: (1) The Parenting Stress Index - Short Form (PSI-SF)

[17]Overwhelming fears which in the context of conflict is associated with long-term complex or post-traumatic stress disorders; see Moya (2018) for a discussion.

[18]Symptoms of negative anger that reflect traits of aggression, irritability, anger, and resentment.

[19]The degree of emotional responses to external stimuli.

a scale that assesses the overall stress of the parenting relationship and captures subconstructs of Parental Distress, Parent-Child Dysfunctional Interaction, and Difficult Child; (2) The Being a Mother scale (BAM), which captures the overall well-being of mothers (or caregivers) and subconstructs of emotional closeness with the child, the experience of the child; and the experience of the adult as a caregiver, which points in the direction of maternal confidence and self-efficacy; and (3) The Child-Parent Relationship scale, which measures the quality of the child-caregiver relationship, including subconstructs of relational conflict, closeness, and dependence of the child.[20] We use the data from these three scales to approximate different dimensions of the child-caregiver relationship including reciprocity and interplay, which is related to the quality or type of the attachment. In addition, these three scales also capture the subjective experience of caregivers, which the observational measures miss, and is important to understand in the context of the intervention.

*Child-Caregiver Interactions*: We administered a survey instrument based on the Colombian Longitudinal Survey (Bernal et al., 2015) that measures the extent to which caregivers interact with their children through different activities, like reading, watching tv, and playing in a park, and positive and negative discipline practices. We differentiate the constructs of child-caregiver relationships and interactions because the former conceptually captures the quality or stress in the child-caregiver and attachment relationship, whereas the latter measures the type or style of child-rearing practices and interactions with the child.

*Child Mental Health*: We administered the Trauma Symptom Checklist for Young Children (TSCYC). a scale that captures different psychopathologies and dimensions of children's mental health, including anger, anxiety, depression, posttraumatic stress, and dissociation. As discussed earlier, this data is reported by the caregiver.

[20] We did not administer observational measures of the interactions between caregivers and their children, like the Stange Situation Test, which is considered the gold standard for the assessment of attachment relationships, because we could not visit families in their homes or create private spaces for the observation of the child-parent dyad interaction.

*Early Childhood Development*: We evaluated children's development through developmentally appropriate assessments that rely on structured play-based tasks.[21] At baseline and the first post-intervention follow-up, we administered the Preschool Self-Regulation Assessment (PSRA), which covers self-regulation in emotional, attentional, and behavioral domains for children aged 30-40 months. As children were 1 year older at the 8-month follow-up, we administered the International Development Early Learning Assessment (IDELA), an assessment for children ages 3 to 6, which covers four developmental domains: motor development, emergent literacy, emergent numeracy, social-emotional development, and executive functions. We complemented these assessments using the Brief Toddler Socioemotional Assessment (BITSEA), a caregiver-reported scale that measures the social-emotional development of children including social-emotional and behavioral problems or delays, and deficits in social-emotional competence. [22]

*Household survey*: At baseline, we also administered a standard household socioeconomic survey, which included a module on household-level exposure to conflict-related violence and displacement previously implemented by Ibáñez and Moya (2010) and Moya (2018). At each follow-up, we administered short questionnaires to capture changes in main socioeconomic dimensions and to assess recent (between-wave) exposure to conflict-related violence.

Our main analysis of the impacts of the program will focus on five composite indices. For this purpose, we pooled together all of the items from all of the scales and assessments in each dimension and estimated the indices following Kling et al. (2007) based on the standardized inverse-covariance weighted average of the different items. This allows us to reduce measurement error in any individual scale or assessment and reduce the dimensions of the analysis. We transformed the signs of all items so that positive signs denote

[21] We also measured children's height and weight, which are proxies of children's health and development. Because we did not expect that the program would affect children's physical development, we only measured anthropometrics at baseline and used them as controls in our econometric analysis.

[22] Although the BITSEA was designed for children 1-3 years of age, and children in our sample were between 2 - 5 years at baseline and 3-6 at the 8-month follow-up, we still find variation in the data and that the scale is internally valid as we discuss in the following section in the following subsection.

positive impacts (improvements) and standardized the indices and scale-specific scores using the mean and standard deviation of the control group at each wave. For child-specific outcomes, we also standardized the measures by children's age (in days) so that scores are comparable. After analyzing the impact of the program on the five composite indices, we conduct an exploratory analysis by unpacking the impact across scale-specific aggregate scores and sub-constructs, which we estimate following the same methodology described above, and at the extensive margin using scale-specific risk thresholds. Since not all scales provide risk thresholds for critical symptoms, and risk thresholds are valid for this specific sample without local adaptations, we define within-sample risk thresholds as a score one standard deviation below the mean.

## 5.3 Psychometrics

We assess the internal validity of the five composite indices by estimating the Cronbach Alphas and the Confirmatory Factor Analysis (CFA), which provide a measure of the indices' internal consistency and of whether the factor structure of the indices captures a latent construct, respectively.[23] As a rule of thumb, a value greater than 0.70 in the Cronbach Alpha indicates an appropriate internal consistency. We found that all indices are internally valid at baseline with alphas that range between 0.71 for the early childhood development index to 0.95 for the caregiver mental health index (see Appendix Table A8, Column 1 ). We also found that the Root Mean Square Error of Approximation (RMSEA) is below 0.05 (0.10) in each of the five indices, indicating a good (adequate) fit in capturing a latent construct (see Appendix Table A8, Column 2).

In Appendix Table A9, we replicate this analysis focusing on the aggregate scores for each scale and the scores across different sub-constructs. In Column 1, we observe that all scales have appropriate internal validity at baseline except for the locus of control ($\alpha$

[23]Psychometric analysis is common and required in psychology but has been largely absent in economics. We raise a point on the importance of this type of analysis in general and more so now that economists are becoming increasingly interested in administering early childhood assessments and mental health questionnaires that have been developed in different contexts and that propose standard scoring criteria.

= 0.66), child-parent interactions questionnaire ($\alpha$ = 0.58), and the BAM ($\alpha$ = 0.55). In Column 2 we report the results from the CFA and observe a good fit in the RMSEA for all of the aggregate scores of the different scales as well as in a large share of the sub-constructs.

The analysis indicates above that our composite indices are internally valid and group together items that capture an underlying construct and that most of the individual scales are also internally valid and measure consistent constructs. For this reason, we analyze the impact of the program through the composite indices that use all of the individual scales, and for robustness, we estimate the effects excluding the scales that have inadequate internal validity.

## 5.4 Covid-19 Disruption

We implemented the program for all cohorts and collected all data for Cohorts 1 and 2 as planned, as well as data at baseline and the 1-month follow-up for Cohorts 3 and 4. See Appendix Figure A6 which illustrates the timeline for the data collection and the program's implementation for each cohort. Unfortunately, the Covid-19 pandemic and the prolonged national lockdown implemented by the Colombian government on March 23, 2020, disrupted in-person data collection for the 8-month follow-up for Cohorts 3 and 4, which were scheduled for March and August 2020, respectively. For this reason, we updated our pre-analysis plan in April 2020, with some uncertainty on whether we would be able to collect the remaining data in person, and specified that if wee couldn't, our analysis of the impact of the program at the 8-month follow-up would rely on the data from Cohorts 1 and 2. This implies a smaller sample size and lower statistical power for the 8-month assessment and also prevents us from observing the impact for cohorts 3 and 4 for which we expected larger effects because of the improvements in the program's dosage and quality over time.

In the amendment to our pre-analysis plan, we also registered a short phone interview,

on a subset of the scales on caregiver anxiety, depression, and parenting stress, that we administered to participants of Cohorts 3 and 4 between April and May 2020. We administered this survey hoping that by assessing participants a few weeks after the lockdown they would still not be severely affected by the pandemic. This would have allowed analyzing the 8-month impacts of the program using the full experimental sample for these three outcomes. However, 5 to 7 weeks into the pandemic, we observed severe hardships including income loss and food insecurity for over 70 percent of the sample which translated into heightened symptoms of anxiety, depression, and parenting stress. Because of the extent of this covariate shock, our main analysis for the 8-month follow-up focuses only on data from Cohorts 1 and 2. See Moya et al. (2021) for the analysis of the mental health consequences of the pandemic using this data.

## 5.5 Attrition

Finally, we analyze the attrition across the survey waves, the degree to which they differ across the different treatment arms, and whether they are explained by baseline characteristics. We first found low rates of attrition in the 1 and 8-month follow-ups, where we surveyed 96 and 91 percent of the subjects who had completed the baseline assessment (see Appendix Table A10). These rates vary slightly across cohorts but are within more than reasonable ranges considering the ongoing conflict and the challenges to locating and contacting study participants outside of the childcare centers.[24]

Furthermore, we find that attrition is not explained by differences in baseline characteristics. In Appendix Table A11 we report the sample mean differences between the full sample at baseline and the sample of non-attritors at the first follow-up. We found find few, if any, statistically significant differences. In addition, in Appendix Table A12 we report the results from the joint-significance test and show that we fail to reject the null

[24]We employed different strategies to minimize attrition including (1) changing the moment of the second follow-up to maximize the percentage of caregivers and children who would still be affiliated with the childcare centers; (2) contacting and collecting all data at the childcare centers; (3) providing a monetary incentive of COP $20,000 (approximately 5 US dollars) at each wave; and (4) reaching out to the childcare center staff and community leaders who helped to contact hard-to-reach subjects.

hypothesis that baseline characteristics jointly predict attrition in the 1-month (F-test = 1.31) and 8-month (F-test = 1.394) followups. For these reasons, we consider attrition to be as good as random and do not adjust our estimates for differential attrition.

# 6 Descriptive Statistics

In Table 2, we report descriptive statistics on caregivers and the index child at baseline. On average, caregivers were 29 years of age, 96 percent were female, and 87 percent were the mothers of the index child. The children assessed in the study had 30 months of age at baseline and were on average the second child in the family. The data also highlights the socioeconomic vulnerability of the caregivers in their study and their families. Caregivers had completed 11.85 years of education on average (slightly below the number of years needed for high school graduation), only 13 percent had formal employment, and the average score in a multidimensional asset index was 1.64 standard deviations below the national mean.

As mentioned in the introduction, the data in the table also highlights the pervasiveness of the conflict, as a majority of the caregivers in the study had directly experienced conflict-related violence or had been forcibly displaced. In addition, 16 and 27 percent of caregivers had at-risk symptoms of anxiety and depression, respectively, which are between 2 to 3 times higher than national averages (Moya, 2018), while 27 and 16 percent of their children had at-risk symptoms for anxiety and depression, respectively, also above national averages.

Table 2: Baseline Descriptives and Balance

| | (1) Total | (2) Control | (3) Treatment | (4) Diff (2)-(3) |
|---|---|---|---|---|
| *Panel A. Child* | | | | |
| Child Age in months | 35.30 | 34.98 | 35.60 | -0.62 |
| | (8.376) | (8.231) | (8.503) | [0.451] |
| Female (=1) | 0.50 | 0.48 | 0.51 | -0.03 |
| | (0.500) | (0.500) | (0.500) | [0.027] |
| Child birth order | 1.90 | 1.88 | 1.91 | -0.03 |
| | (1.182) | (1.171) | (1.193) | [0.064] |
| *Panel B. Caregiver* | | | | |
| Caregiver is the mother (=1) | 0.87 | 0.86 | 0.89 | -0.03 |
| | (0.335) | (0.351) | (0.319) | [0.018] |
| Female (=1) | 0.96 | 0.94 | 0.97 | -0.03** |
| | (0.199) | (0.236) | (0.157) | [0.011] |
| Age | 29.06 | 28.89 | 29.22 | -0.33 |
| | (9.259) | (9.015) | (9.484) | [0.499] |
| Years of education | 11.85 | 11.86 | 11.84 | 0.02 |
| | (3.673) | (3.552) | (3.785) | [0.198] |
| *Panel C. Household* | | | | |
| Household size | 5.00 | 4.98 | 5.02 | -0.04 |
| | (1.992) | (1.934) | (2.045) | [0.107] |
| Number of children under 5yrs | 1.29 | 1.26 | 1.32 | -0.06 |
| | (0.558) | (0.495) | (0.610) | [0.030] |
| Two-parent hh (=1) | 0.70 | 0.68 | 0.73 | -0.05* |
| | (0.457) | (0.468) | (0.446) | [0.025] |
| Highest years of education | 12.69 | 12.54 | 12.83 | -0.29 |
| | (3.314) | (3.315) | (3.309) | [0.179] |
| Asset index | -1.64 | -1.70 | -1.59 | -0.11 |
| | (1.197) | (1.189) | (1.202) | [0.064] |
| Hh income per capita | 242.86 | 249.89 | 236.35 | 13.53 |
| | (316.613) | (339.105) | (294.319) | [17.176] |
| CCT beneficiary (=1) | 0.43 | 0.42 | 0.43 | -0.02 |
| | (0.495) | (0.493) | (0.496) | [0.027] |
| Head is employed (=1) | 0.83 | 0.83 | 0.83 | -0.00 |
| | (0.379) | (0.380) | (0.378) | [0.020] |
| Head has formal job (=1) | 0.21 | 0.19 | 0.22 | -0.03 |
| | (0.406) | (0.393) | (0.417) | [0.022] |
| Days hh head works in a month | 23.75 | 23.58 | 23.92 | -0.34 |
| | (6.429) | (6.534) | (6.334) | [0.388] |
| *Panel D. Violence* | | | | |
| Victim of direct violence (=1) | 0.82 | 0.80 | 0.84 | -0.04* |
| | (0.385) | (0.403) | (0.367) | [0.021] |
| Number of violent events | 2.36 | 2.23 | 2.48 | -0.25** |
| | (1.799) | (1.771) | (1.817) | [0.097] |
| IDP (=1) | 0.57 | 0.56 | 0.58 | -0.02 |
| | (0.495) | (0.496) | (0.494) | [0.027] |
| *Panel E. Indices* | | | | |
| Caregiver Mental Health | -0.06 | -0.00 | -0.11 | 0.11* |
| | (0.800) | (0.762) | (0.832) | [0.043] |
| Child-Caregiver Relationships | -0.03 | -0.00 | -0.06 | 0.06 |
| | (0.847) | (0.820) | (0.872) | [0.046] |
| Child Mental Health | -0.06 | 0.00 | -0.11 | 0.11 |
| | (1.118) | (1.000) | (1.215) | [0.060] |
| Child Development | -0.03 | 0.00 | -0.05 | 0.05 |
| | (0.792) | (0.742) | (0.837) | [0.043] |
| Observations | 1376 | 662 | 714 | 1376 |
| Joint F-statistic | 0.649 | | | |

Notes:

# 7 Results

In this section, we report the results of the program taking advantage of the random assignment at the childcare center level. We start by estimating the effects of the program on the composite index for each of the five dimensions of interest of the program: caregiver mental health, child-caregiver relationships, child-caregiver interactions, child mental health, and early childhood development. Then we conduct an exploratory analysis focused on the impact on the individual scales, both at the intensive margin —the continuous z-scores provided by each scale— and at the extensive margin, the likelihood of at-risk symptoms. Finally, we report results from the heterogeneity analyses.

To complement these results, and as a strategy to overcome the limitations in power from the cluster randomization, we exploit the phase-in design of the program for robustness.

## 7.1 Empirical Strategy

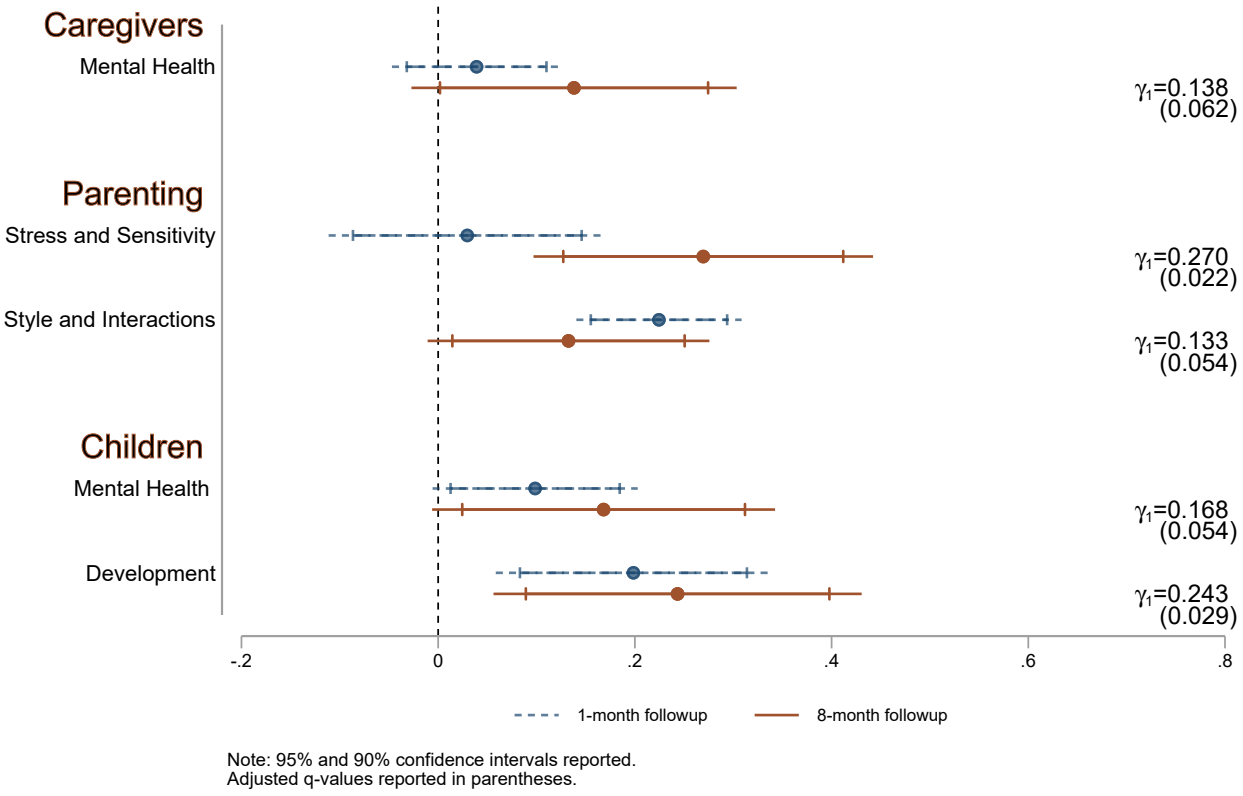
In our main analyses, we estimate the causal impact of *Semillas de Apego* by exploiting the random assignment at the childcare center level. We estimate a lagged-dependent variable model where we regress the outcome $Y_i^f$ at the 1 or 8-month followup $f$, on the treatment assignment $Treatment_i$, the outcome at baseline $Y_i^f$, a matrix $X^{BL}$ of baseline individual and household controls, and a set of cohort fixed effects $\phi_c$. Since we did not pre-register the set of covariates, we selected those that are strongly predictive of the outcome (regardless of the balance) and we will report results with and without adjusted specifications. We cluster our standard errors at the childcare center level and follow Pustejovsky and Tipton (2018b) to account for the small number of clusters. We also adjust our estimates for multiple hypothesis testing using Anderson's q-value (Anderson, 2008).

$$Y_i^f = \gamma_0 + \gamma_1 Treatment_i + \gamma_2 Y_i^{bl} + \Gamma' X^{bl} + \phi_c + \epsilon_i$$

## 7.2 Main Results

In Figure 3, we provide a graphical summary of the main results. In the figure, we plot the point estimates and 95 and 90th percentile confidence intervals for the program's impact on the five main indices at the two follow-ups. Overall, the figure indicates that at the 1-month follow-up, the program had sizeable and statistically significant impacts on children's mental health and development but, surprisingly, without any discernable effects on caregivers' mental health or the quality of the child-caregiver relationship. At the 8-month follow-up, we observe meaningful and statistically significant effects in these two latter dimensions and stronger impacts on children's outcomes than at the first assessment. Taken together, the results suggest that the program's impacts build up over time, which is important given the evidence in the ECD space on impacts that fade out in time. Below, we analyze these results in detail.

Figure 3: Impact at 1 and 8-month followups




Note: 95% and 90% confidence intervals reported.
Adjusted q-values reported in parentheses.

## 7.3 1-month impacts

In Table 3, we report the results of the program's impact on the composite indices at the 1 and 8-month follow-ups. In the table, each panel includes the coefficients of interest for each of the five composite indices that result from estimating model 1 under the full specification –this is, after controlling for the lagged dependent variable, household and individual controls, and cohort fixed effects.[25] We report the ITT and the Average Treatment on the Treated (ATT) at each follow-up, where the latter is estimated using the random assignment into the program to instrument for the participation in at least 8 sessions, the unconditional average dosage. For each point estimate, we report the CR2 standard error that corrects for the small number of clusters and Andersen's q-value that adjusts for multiple hypothesis testing.

At the first follow-up, we observed that the program had positive and statistically significant impacts on child-caregiver interactions, child mental health, and early childhood development with effects ranging from 0.1 to 0.23 (see Column 1). For caregivers who participated in 8 or more sessions, the ATT's are also statistically significant and large, with effect sizes ranging between 0.16 and 0.37 sd (see Column 2), which is explained by the fact that no caregivers in the control arm had access to the program and that 65 percent of caregivers assigned to the treatment arm attended more than 8 sessions. Overall, the point estimates for the three composite indices are consistent across the different specifications and illustrate that we gain precision and power by controlling for the baseline prognostic variables.

We also find that the program had small and not-different from zero impacts on caregiver mental health and child-caregiver relationships, which are our proxy of the quality of the relationship and emotional attachment. The point estimates in these two dimensions are small and close to zero, pointing to an imprecisely estimated null effect.

[25] For transparency, in the Appendix we report the results of model 1 when we only control for the baseline outcome and then as we sequentially control the cohort fixed effects and baseline individual and household controls.

Table 3: Results at 1-month and 8-month follow-up

| | 1-month | | 8-month | |
|---|---|---|---|---|
| | ITT | ATT | ITT | ATT |
| | (1) | (2) | (3) | (4) |
| *Panel A. Caregiver Mental Health* | | | | |
| Treatment | 0.039 | 0.065 | 0.138* | 0.264* |
| | [0.041] | [0.066] | [0.078] | [0.147] |
| | (0.416) | (0.194) | (0.062) | (0.046) |
| Observations | 1317 | 1317 | 522 | 522 |
| *Panel B. Child-Parent Relationships* | | | | |
| Treatment | 0.030 | 0.049 | 0.270*** | 0.520*** |
| | [0.067] | [0.108] | [0.082] | [0.165] |
| | (0.475) | (0.349) | (0.022) | (0.009) |
| Observations | 1318 | 1318 | 522 | 522 |
| *Panel C. Child-Parent Interactions* | | | | |
| Treatment | 0.225*** | 0.373*** | 0.133* | 0.256** |
| | [0.040] | [0.064] | [0.068] | [0.129] |
| | (0.001) | (0.001) | (0.054) | (0.037) |
| Observations | 1317 | 1317 | 522 | 522 |
| *Panel D. Early Childhood Mental Health* | | | | |
| Treatment | 0.099* | 0.164** | 0.168* | 0.329** |
| | [0.049] | [0.081] | [0.083] | [0.158] |
| | (0.214) | (0.044) | (0.054) | (0.037) |
| Observations | 1315 | 1315 | 518 | 518 |
| *Panel E. Early Childhood Development* | | | | |
| Treatment | 0.199*** | 0.352*** | 0.243** | 0.480*** |
| | [0.066] | [0.117] | [0.089] | [0.183] |
| | (0.004) | (0.006) | (0.034) | (0.029) |
| Observations | 937 | 937 | 504 | 504 |

Notes: Results of the lagged dependent variable model where the outcome indicated in each panel is measured at the 1-month and 8-month follow-up respectively is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of household controls in the baseline. Controls for Panel A include caregiver's age, asset index, two-parent household, number of conflict-related violent events, whether the participant has been internally displaced, whether the household head had a formal job, and whether the participant had worked in the previous 7 days. Controls for Panel B additionally include the child's sex, the child's birth order, and the caregiver's mental health. Finally, controls for Panels C and D additionally include the child's weight for height Z-score. Columns 1 and 3 report the Intent-to-Treat effects, while Columns 2 and 4 report the Average Treatment on the Treated effect for attending the average number of sessions following a two-stage least squares model. Cluster robust standard errors are reported in brackets (CR2).; * $\rho$ <0.1. ** $\rho$ <0.05. *** $\rho$ <0.01.

Furthermore, the results are robust when we re-estimate our composite indices excluding scales or sub-scales with sub-standard psychometric properties.

At face value, the results in Columns 1 and 2 of Table 3 seem at odds with the underlying hypothesis that motivates the *Semillas de Apego* program. In particular, if caregivers' mental health is a binding constraint and healthy attachment relationships are required to protect young children from the effects of conflict and forced displacement, it is perhaps surprising that the program had no effect on caregivers' mental health and the quality of the child-caregiver relationship but did promote children's mental health and development. However, this result is consistent with clinical studies of the *Child-Parent Psychotherapy*, which provided the conceptual and implementation basis for *Semillas de Apego* and the pilot study of the program, where we observed a null immediate impact on caregivers' mental health. While speculative, this can be explained by the fact that the program places a large focus on allowing caregivers to be better aware of and attuned to their mental health and the quality of their relationship with their children. As a result, caregivers in the treatment arm may report more symptoms of mental health problems as they become more aware relative to caregivers in the control arm, and the estimated impact may be masking the improvements in mental health with this greater awareness. Similar effects have been observed in psycho-education programs.

## 7.4 8-month impacts

In Columns 3 and 4 of Table 3, we report the results at the 8-month follow-up. Although this analysis is based on a smaller sample (participants in Cohorts 1 and 2) because of the disruption from the COVID-19 pandemic, and confidence intervals are larger as a result, we find positive impacts on all five dimensions. Specifically, we find ITT of 0.14 standard deviations on the maternal mental health index; 0.27 sd on the child-caregiver relationships; 0.13 sd on the style of interactions; and 0.17 and 0.24 sd on children's mental health and early-childhood development, respectively. Likewise, ATT estimates are

stronger (as expected) and statistically significant across the five domains. These range from 0.26 sd on caregivers' mental health to 0.52 and 0.48 sd effects on the quality of the child-caregiver relationships and children's early childhood development. The latter captures a sizeable and meaningful effect, larger than those of standard parenting interventions in the region. As before, these results are consistent in the more parsimonious specifications and as we sequentially control for the lagged (baseline) dependent variable, household and individual fixed effects, and cohorts fixed effects.

Taken together, the results above indicate that the program had impacts that grew and persisted over time, consistent with the conceptual framework and theory of change. Further work, however, is needed to identify the mechanisms of impact and to test whether the effects on children's outcomes are mediated by the improvements in caregiver's capacities and mental health. Yet, the effects are illustrative of the intervention generating sustained impacts on caregivers' emotional regulation as a stepping stone to promote healthier relationships that promote adequate emotional regulation in children, even in the context of protracted conflict and ongoing trauma.

## 7.5 Unpacking the composite indices

To better understand the impact of the program and the possible mechanisms of impact, we now unpack the composite indices and report the results on the individual scales or subconstructs. We conduct this analysis at the intensive and extensive margin; e.g., for both the continuous scores and the likelihood of scoring above the risk thresholds. Because of the number of tests we conduct, we consider this analysis as exploratory and do not adjust for multiple hypothesis testing.

In Table 4, we report the impacts across subdimensions at the 1 and 8-month follow-ups focusing on the intensive margin. In general, we find impacts that are consistent with those of the composite indices, suggesting that the program generated improvements in the type of interactions between caregivers and their children, including a higher number of

interactions every week and more interactions characterized by positive discipline strategies and less with negative discipline or harsh parenting (including verbal and physical violence). Likewise, we observe that the impact on early childhood development is driven by (self-reported) improvements of 0.14sd in children's socioemotional competencies. We observe a positive and sizeable effect of 0.32 sd on children's development assessment that is not statistically different from zero.

Table 4: Results by subdimensions: 1-month and 8-month followup

| | 1-month | | 8-month | |
|---|---|---|---|---|
| | ITT | TOT | ITT | TOT |
| *Panel A. Caregiver Mental Health* | | | | |
| Anxiety | -0.008 | -0.012 | 0.056 | 0.148 |
| | [0.042] | [0.065] | [0.070] | [0.138] |
| Depression | 0.013 | 0.021 | 0.019 | 0.028 |
| | [0.039] | [0.060] | [0.081] | [0.149] |
| Phobic Anxiety | 0.004 | 0.007 | 0.123 | 0.279* |
| | [0.046] | [0.071] | [0.074] | [0.148] |
| Sensitivity | -0.036 | -0.058 | 0.170** | 0.290** |
| | [0.049] | [0.075] | [0.079] | [0.145] |
| Hostility | 0.012 | 0.019 | 0.203*** | 0.329*** |
| | [0.042] | [0.064] | [0.057] | [0.124] |
| IPPA | | | -0.027 | -0.065 |
| | | | [0.062] | [0.134] |
| Internal Locus of Control | 0.048 | 0.076 | 0.068 | 0.149 |
| | [0.038] | [0.059] | [0.076] | [0.118] |
| *Panel B. Child-Caregiver Relationships* | | | | |
| Parenting Stress Index (PSI) | 0.023 | 0.037 | 0.187* | 0.344* |
| | [0.074] | [0.114] | [0.100] | [0.179] |
| Being a Mother | -0.022 | -0.035 | 0.137* | 0.252** |
| | [0.074] | [0.113] | [0.069] | [0.127] |
| Child-Parent Relationship Scale | 0.066 | 0.106 | 0.299*** | 0.551*** |
| | [0.053] | [0.083] | [0.084] | [0.157] |
| *Panel C. Child-Parent Interactions* | | | | |
| Interactions | 0.209*** | 0.334*** | 0.110 | 0.203* |
| | [0.035] | [0.056] | [0.065] | [0.117] |
| Positive discipline | 0.147** | 0.235** | 0.075 | 0.140 |
| | [0.060] | [0.094] | [0.078] | [0.140] |
| Negative discipline | 0.112** | 0.178** | 0.106 | 0.199 |
| | [0.047] | [0.072] | [0.093] | [0.164] |
| Routines | 0.111** | 0.177** | 0.006 | 0.012 |
| | [0.045] | [0.069] | [0.087] | [0.155] |
| *Panel D. Child Mental Health* | | | | |
| Anger | 0.079 | 0.125* | 0.135 | 0.250 |
| | [0.049] | [0.076] | [0.087] | [0.153] |
| Anxiety | 0.121** | 0.194*** | 0.110 | 0.205 |
| | [0.043] | [0.066] | [0.086] | [0.152] |
| Depression | 0.069 | 0.110 | 0.251** | 0.469*** |
| | [0.057] | [0.088] | [0.095] | [0.174] |
| Dissociation | 0.018 | 0.028 | 0.090 | 0.167* |
| | [0.060] | [0.093] | [0.054] | [0.093] |
| Post Traumatic Stress | 0.019 | 0.030 | 0.139 | 0.258 |
| | [0.063] | [0.096] | [0.116] | [0.208] |
| Sex Concern | -0.044 | -0.070 | 0.062 | 0.115 |
| | [0.051] | [0.077] | [0.074] | [0.131] |
| *Panel E. Early Childhood Development* | | | | |
| Assessments | 0.318 | 0.538 | 0.510 | 0.859 |
| | [0.288] | [0.458] | [1.530] | [2.749] |
| Bitsea | 0.147*** | 0.235*** | 0.227** | 0.423** |
| | [0.044] | [0.070] | [0.089] | [0.168] |
| Observations | 1315 | 1315 | 518 | 518 |
| Controls in Baseline | Yes | Yes | Yes | Yes |
| Cohort Fixed Effects | Yes | Yes | Yes | Yes |

Differently, at the 8-month follow-up, we estimate sizeable and statistically significant effects across caregivers' mental health and children's mental health dimensions. Again, the estimates at the 8-month follow-up are less precise because of the smaller sample size used in this analysis. Nevertheless, the point estimates suggest stronger effects in the 8-month follow-up than in the 1-month follow-up, consistent with the results discussed earlier on the composite indices. For caregiver's mental health, for example, we observe improvements (reductions of symptoms) of 0.17 and 0.20 sd in sensitivity and hostility, which also emerged during the program's pilot and are consistent with the program's design and sessions that attempt to create greater emotional regulation. By contrast, we find small and null impacts on anxiety and depression, which are the usual dimensions that prior studies in economics have focused on but that, as the results suggest, may miss impacts across other dimensions or areas of mental health. Furthermore, we find larger effects on caregiver-self-reported assessment of children's socioemotional competencies (a 0.22 sd ITT effect) and now we also find sizeable improvements (of 0.25 sd) on children's symptoms of depression, which seem to drive the effect of the composite index on children's mental health.

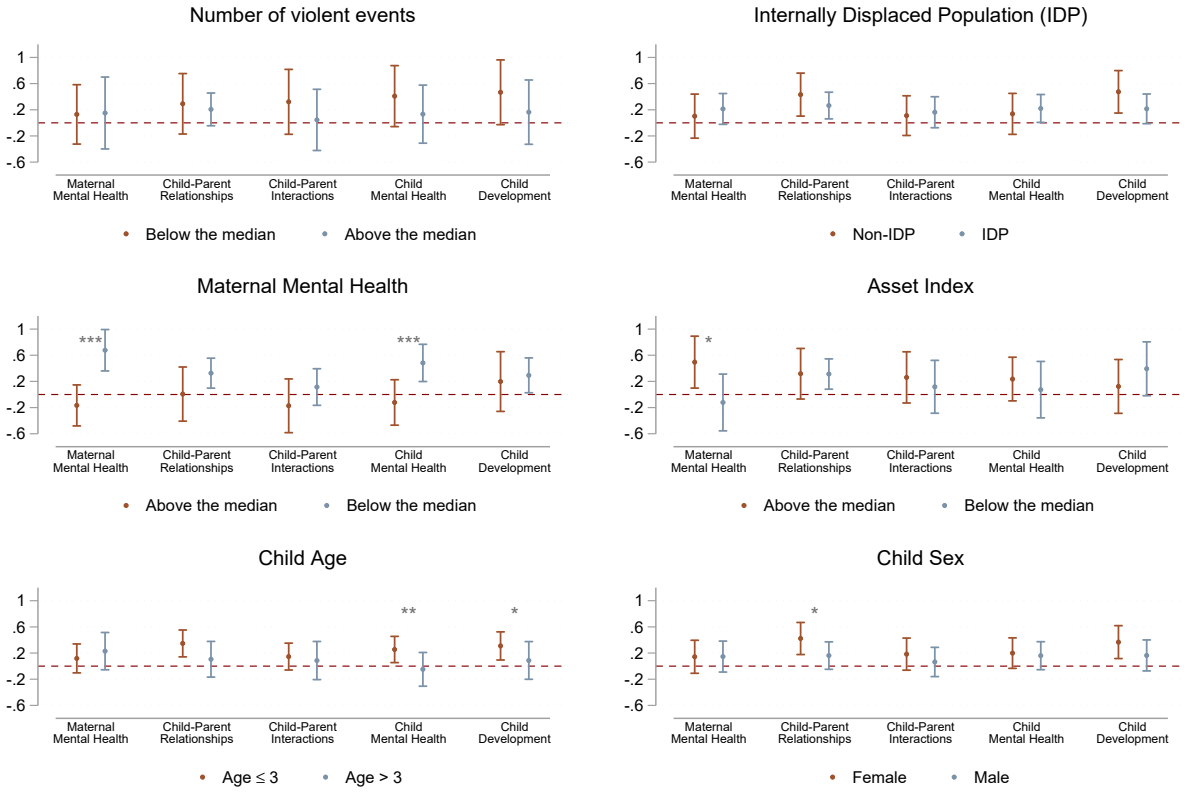
Finally, in Table 4 we analyze the impacts on the likelihood of at-risk symptoms. Different from what we observed in the analysis of the composite indices, at the 1-month follow-up we observe effects across many of the sub dimensions indicating that the program was effective at the extensive margin; in reducing the likelihood of at-risk symptoms. Specifically, at the 1-month follow-up we see a statistically significant reduction in the at-risk likelihood for different dimensions of maternal mental health, for the parenting stress index and the battery of child-parent interactions, and for the different dimensions of child mental health. For early childhood development, we only see a reduction in the at-risk symptoms following the caregiver-reported BITSEA for socioemotional competence but not on the preschool readiness assessment, which is a more objective measure of early childhood development.

## 7.6 Heterogeneity of impact

To assess heterogeneity in impact, we followed two different strategies. First, we estimate model 1 above stratifying the data according to baseline covariates that point to different dimensions of vulnerability: conflict-exposure, mental health, and socioeconomic wellbeing. For example, we estimate model 1 above separately for IDPs and non-IDPs and then test for whether the point estimates for the effect of the treatment are statistically different. Second, we implement random forest algorithms proposed by Wager and Athey (2018) that use machine-learning techniques to identify key dimensions of heterogeneity. We take these two analyses as exploratory because uncovering heterogeneous effects requires larger sample sizes and our sample may be underpowered.

Figure 4 summarizes the results of the parsimonious approach. We find heterogeneity of impacts across three main dimensions: exposure to prior violence and pre-existing mental health problems, and socioeconomic vulnerability (see Appendix Tables A15-A18). On the one hand, we find that the program has stronger effects for IDPs and for those with above-average symptoms of mental health problems at baseline. On the other hand, we find that the program does not have sizeable effects on those who were food-insecure and had baseline levels of wealth below the sample median.

Figure 4: Heterogeneous treatment effects




Statistical difference: * p < .01, ** p < .05, *** p < .1

## 7.7 Robustness: Within-treatment variation.

Finally, we take advantage of the step-wedge design and exploit random assignment to the different cohorts within the treated childcare centers. This strategy allows us to provide an alternative take on the results that is not based on the randomization of the program across a small number of clusters.

We report the results on the four composite indices below in Table 5. The results from this analysis by and large confirm the results that are based on the random assignment at the childcare centers. Specifically, we observe sizeable and statistically significant ITT and TOT effects of the program on the four main dimensions at the 8-month followup.

Table 5: Results: Within-Treatment (ITT)

| | ITT | |
|---|---|---|
| | 1-month followup | 8-month followup |
| *Panel A. Caregiver Mental Health* | | |
| Treatment | 0.222*** | 0.268*** |
| | [0.079] | [0.086] |
| | (0.011) | (0.002) |
| Observations | 702 | 689 |
| *Panel B. Child-Parent Relationships* | | |
| Treatment | 0.075 | 0.378*** |
| | [0.071] | [0.079] |
| | (0.196) | (0.001) |
| Observations | 703 | 689 |
| *Panel B. Child-Parent Interactions* | | |
| Treatment | 0.381*** | 0.160** |
| | [0.068] | [0.076] |
| | (0.001) | (0.008) |
| Observations | 702 | 689 |
| *Panel C. Child Mental Health* | | |
| Treatment | 0.078 | 0.266*** |
| | [0.077] | [0.080] |
| | (0.196) | (0.001) |
| Observations | 700 | 685 |
| *Panel D. Child Development* | | |
| Treatment | 0.124* | 0.348*** |
| | [0.075] | [0.081] |
| | (0.109) | (0.001) |
| Observations | 687 | 679 |
| Cohort Fixed Effects | Yes | Yes |
| Controls | Yes | Yes |

Notes: Results of the lagged dependent variable model exploiting the within-treatment random assignment of the program across cohorts. The outcome indicated in each panel and measured at the 1-month and 8-month followup respectively is regressed on treatment assignment, the lagged-dependent variable at baseline, and a set of household controls in Baseline. Controls are the same as those in prior tables. Robust standard errors are reported in brackets. * $\rho$ <0.1. ** $\rho$ <0.05. *** $\rho$ <0.01.

# 8 Discussion

We provide a compelling argument for the importance of early childhood development and the impact of conflict and forced migration on children's well-being. Our paper highlights the critical role of adult caregivers in providing nurturing care for children in settings of extreme adversity and trauma, and the need for interventions to promote maternal mental health as a pathway to protect children's development. Our focus on a community and group-based psychosocial model to support caregivers of young children affected by conflict or forced migration represents an innovative approach that addresses the maternal mental health constraint in such settings.

We conducted the study in a conflict-torn municipality in Colombia, which has been historically affected by conflict and forced displacement and is an appropriate context to evaluate the intervention's effectiveness. We targeted mothers or other primary caregivers of children aged 2 to 5 served by public childcare centers serving vulnerable and underserved families. The experimental evaluation design overcame the challenges brought forth by the ongoing conflict and the resource-constrained setting by randomizing the assignment of treatment at the childcare center level, implementing the model across four sequential cohorts, and collecting all data at the childcare centers to ensure the safety of participants, enumerators, and of the implementation and research teams.

Our results suggest that the community and group-based psychosocial model positively impacted maternal mental health, nurturing care practices, and children's developmental outcomes. More broadly speaking, our findings speak to the need and feasibility of implementing quality psychosocial programs in fragile and conflict-affected settings but also on the importance of designing comprehensive strategies that address social and economic determinants of mental health.

For further discussion, we will explore the implications for policy and practice and the potential for scaling up the intervention to other contexts affected by conflict and forced migration.

# References


**Anderson, Michael L**, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495. 5, 31

**Andrew, Alison, Orazio Attanasio, Britta Augsburg, Monimalika Day, Sally Grantham-McGregor, Costas Meghir, Fardina Mehrin, Smriti Pahwa, and Marta Rubio-Codina**, "Effects of a scalable home-visiting intervention on child development in slums of urban India: evidence from a randomised controlled trial," *Journal of child psychology and psychiatry*, 2020, *61* (6), 644–652. 6

**Angelucci, Manuela and Daniel Bennett**, "The economic impact of depression treatment in india: Evidence from community-based provision of pharmacotherapy," *American Economic Review*, 2024, *114* (1), 169–198. 8

**Araujo, M Caridad, Marta Dormal, Sally Grantham-McGregor, Fabiola Lazarte, Marta Rubio-Codina, and Norbert Schady**, "Home visiting at scale and child development," *Journal of Public Economics Plus*, 2021, *2*, 100003. 6

**Ariza, Juliana Sánchez, Jorge Cuartas, and Andrés Moya**, "The Mental Health of Caregivers and Young Children in Conflict-Affected Settings," *American Economic Association Papers and Proceedings*, 2023, *X* (X), X. 2, 7

**Attanasio, Orazio, Helen Baker-Henningham, Raquel Bernal, Costas Meghir, Diana Pineda, and Marta Rubio-Codina**, "Early Stimulation and Nutrition: the impacts of a scalable intervention," *Journal of the European Economic Association*, 2022, *20* (4), 1395–1432. 6

**_ , Raquel Bernal, Michele Giannola, and Milagros Nores**, "Child development in the early years: Parental investment and the changing dynamics of different dimensions," 2020. 18

**_ , Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina**, "Estimating the production function for human capital: results from a randomized controlled trial in Colombia," *American Economic Review*, 2020, *110* (1), 48–85. 6

**Baranov, Victoria, Sonia Bhalotra, Pietro Biroli, and Joanna Maselko**, "Maternal depression, women's empowerment, and parental investment: evidence from a randomized controlled trial," *American economic review*, 2020, *110* (3), 824–59. 8

**Bernal, Raquel and Sara María Ramírez**, "Improving the quality of early childhood care at scale: The effects of "From Zero to Forever"," *World Development*, 2019, *118*, 91–105. 15

**_ , María Adelaida Martínez, and Claudia Quintero Salleg**, "Situación de niñas y niños colombianos menores de cinco años, entre 2010 y 2013," Technical Report, Universidad de los Andes-CEDE 2015. 24, A11

_ , **Orazio Attanasio, Ximena Peña, and Marcos Vera-Hernández**, "The effects of the transition from home-based childcare to childcare centers on children's health and development in Colombia," *Early Childhood Research Quarterly*, 2019, *47*, 418–431. 15

**Black, Maureen M, Susan P Walker, Lia CH Fernald, Christopher T Andersen, Ann M DiGirolamo, Chunling Lu, Dana C McCoy, Günther Fink, Yusra R Shawar, Jeremy Shiffman et al.**, "Advancing Early Childhood Development: From Science to Scale 1: Early childhood development coming of age: Science through the life course," *Lancet (London, England)*, 2017, *389* (10064), 77. 2

**Blattman, Christopher and Edward Miguel**, "Civil war," *Journal of Economic literature*, 2010, *48* (1), 3–57. 8

**Bossuroy, Thomas, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C Thomas, Christopher Udry, Julia Vaillant et al.**, "Tackling psychosocial and capital constraints to alleviate poverty," *Nature*, 2022, *605* (7909), 291–297. 8

**Briere, John, Kerri Johnson, Angela Bissada, Linda Damon, Julie Crouch, Eliana Gil, Rochelle Hanson, and Vickie Ernst**, "The Trauma Symptom Checklist for Young Children (TSCYC): reliability and association with abuse exposure in a multi-site study," *Child Abuse & Neglect*, 2001, *25*, 1001–1014. A11

**Briggs-Gowan, Margaret J and Alice S Carter**, "Brief infant toddler social emotional assessment," 2007. A11

**Britto, Pia R, Stephen J Lye, Kerrie Proulx, Aisha K Yousafzai, Stephen G Matthews, Tyler Vaivada, Rafael Perez-Escamilla, Nirmala Rao, Patrick Ip, Lia CH Fernald et al.**, "Nurturing care: promoting early childhood development," *The Lancet*, 2017, *389* (10064), 91–102. 2

**Camacho, Adriana**, "Stress and birth weight: evidence from terrorist attacks," *American Economic Review*, 2008, *98* (2), 511–515. 7

**Currie, Janet and Douglas Almond**, "Human capital development before age five," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1315–1486. 2

**Derogatis, Leonard R and Rachael Unger**, "Symptom checklist-90-revised," *The Corsini encyclopedia of psychology*, 2010, pp. 1–2. A11

**Driscoll, Kate and Robert C Pianta**, "Mothers' and Fathers' Perceptions of Conflict and Closeness in Parent-Child Relationships during Early Childhood.," *Journal of Early Childhood & Infant Psychology*, 2011, (7). A11

**Emmers, Dorien, Juan Carlos Caro, Scott Rozelle, and Sean Sylvia**, "Early Parenting Interventions to Foster Human Capital in Developing Countries," *Annual Review of Resource Economics*, 2022, *14*, 169–192. 6

**Evans, David K, Pamela Jakiela, and Heather A Knauer**, "The impact of early childhood interventions on mothers," *Science*, 2021, *372* (6544), 794–796. 7

**Fontes, L F, M Mrejen, B Rache, and R Rocha**, "Economic Distress and Children's Mental Health: Evidence from the Brazilian High Risk Cohort Study for Mental Conditions*," *The Economic Journal*, 12 2023, p. uead109. 7

**Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M Chang, and Sally Grantham-McGregor**, "Labor market returns to an early childhood stimulation intervention in Jamaica," *Science*, 2014, *344* (6187), 998–1001. 6

**Grantham-McGregor, Sally, Akanksha Adya, Orazio Attanasio, Britta Augsburg, Jere Behrman, Bet Caeyers, Monimalika Day, Pamela Jervis, Reema Kochar, Prerna Makkar et al.**, "Group sessions or home visits for early childhood development in India: a cluster RCT," *Pediatrics*, 2020, *146* (6). 6

**Grantham-McGregor, Sally M, Christine A Powell, Susan P Walker, and John H Himes**, "Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study," *The Lancet*, 1991, *338* (8758), 1–5. 6

**Grupo de Memoria Histórica**, *¡ Basta ya! Colombia: Memorias de guerra y dignidad (bogotá: Imprenta Nacional), 431 pp. 1* 2014. 14

**Haskett, Mary E, Lisa S Ahern, Caryn S Ward, and Jason C Allaire**, "Factor structure and validity of the parenting stress index-short form," *Journal of Clinical Child & Adolescent Psychology*, 2006, *35* (2), 302–312. A11

**Haushofer, Johannes and Ernst Fehr**, "On the psychology of poverty," *science*, 2014, *344* (6186), 862–867. 7

**Ibáñez, Ana María and Andrés Moya**, "Vulnerability of victims of civil conflicts: empirical evidence for the displaced population in Colombia," *World development*, 2010, *38* (4), 647–663. 25

**Jeong, Joshua, Emily E Franchett, Clariana V Ramos de Oliveira, Karima Rehmani, and Aisha K Yousafzai**, "Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis," *PLoS medicine*, 2021, *18* (5), e1003602. 7

**Kass, Jared D, Richard Friedman, Jane Leserman, Margaret Caudill, Patricia C Zuttermeister, and Herbert Benson**, "An inventory of positive psychological attitudes with potential relevance to health outcomes: Validation and preliminary testing," *Behavioral Medicine*, 1991, *17* (3), 121–129. A11

**Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz**, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119. 5, 25

**Knudsen, Eric I, James J Heckman, Judy L Cameron, and Jack P Shonkoff**, "Economic, neurobiological, and behavioral perspectives on building America's future workforce," *Proceedings of the national Academy of Sciences*, 2006, *103* (27), 10155–10162. 2

**Lieberman, Alicia F and Patricia Van Horn**, *Psychotherapy with infants and young children: Repairing the effects of stress and trauma on early attachment*, Guilford Press, 2011. 3, 12

**Luo, Renfu, Dorien Emmers, Nele Warrinnier, Scott Rozelle, and Sean Sylvia**, "Using community health workers to deliver a scalable integrated parenting program in rural China: A cluster-randomized controlled trial," *Social Science & Medicine*, 2019, *239*, 112545. 6

**Maslow, Abraham Harold**, "A theory of human motivation.," *Psychological review*, 1943, *50* (4), 370. 6

**Matthey, Stephen**, "Assessing the experience of motherhood: the Being a Mother Scale (BaM-13)," *Journal of affective disorders*, 2011, *128* (1-2), 142–152. A11

**McLaughlin, Katie A, David Weissman, and Debbie Bitrán**, "Childhood adversity and neural development: A systematic review," *Annual review of developmental psychology*, 2019, *1*, 277–312. 2

**Mollica, Richard F, Keith McInnes, Charles Pool, and Svang Tor**, "Dose-effect relationships of trauma to symptoms of depression and post-traumatic stress disorder among Cambodian survivors of mass violence," *The British journal of psychiatry*, 1998, *173* (6), 482–488. 9

**Moya, Andrés**, "Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in Colombia," *Journal of Development Economics*, 2018, *131*, 15–27. 7, 9, 23, 25, 29

**_ , Pieter Serneels, Alethea Desrosiers, Vilma Reyes, María José Torres, and Alicia Lieberman**, "The COVID-19 pandemic and maternal mental health in a fragile and conflict-affected setting in Tumaco, Colombia: a cohort study," *The Lancet Global Health*, 2021, *9* (8), e1068–e1076. 28

**Murphy, Katie Maeve, Hirokazu Yoshikawa, and Alice J Wuermli**, "Implementation research for early childhood development programming in humanitarian contexts," *Annals of the New York Academy of Sciences*, 2018, *1419* (1), 90–101. 8

**Nelson, Charles A and Laurel J Gabard-Durnam**, "Early adversity and critical periods: neurodevelopmental consequences of violating the expectable environment," *Trends in neurosciences*, 2020, *43* (3), 133–143. 2

**Persson, Petra and Maya Rossin-Slater**, "Family ruptures, stress, and the mental health of the next generation," *American economic review*, 2018, *108* (4-5), 1214–52. 7

**Pisani, Lauren, Ivelina Borisova, and Amy Jo Dowd**, "Developing and validating the international development and early learning assessment (IDELA)," *International Journal of Educational Research*, 2018, *91*, 1–15. A11

**Pustejovsky, James E and Elizabeth Tipton**, "Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models," *Journal of Business & Economic Statistics*, 2018, *36* (4), 672–683. 5

**Pustejovsky, James E. and Elizabeth Tipton**, "Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models," *Journal of Business & Economic Statistics*, 2018, *36* (4), 672–683. 31

**Ridley, Matthew, Gautam Rao, Frank Schilbach, and Vikram Patel**, "Poverty, depression, and anxiety: Causal evidence and mechanisms," *Science*, 2020, *370* (6522), eaay0214. 7

**Roa, Arturo Harker, Natalia Córdoba Flechas, Andrés Moya Rodríguez, and Maria Fernanda Pineros Leano**, "Implementing psychosocial support models in contexts of extreme adversity: Lessons from a process evaluation in Colombia," *Frontiers in Psychology*, 2023, *14*, 1282. 19

**Rotter, Julian B**, "Generalized expectancies for internal versus external control of reinforcement.," *Psychological monographs: General and applied*, 1966, *80* (1), 1. A11

**Sánchez-Ariza, Juliana, Jorge Cuartas, and Andrés Moya**, "The Mental Health of Caregivers and Young Children in Conflict-Affected Settings," in "AEA Papers and Proceedings," Vol. 113 2023, pp. 336–41. 11

**Shonkoff, Jack P, Andrew S Garner, Benjamin S Siegel, Mary I Dobbins, Marian F Earls, Laura McGuinn, John Pascoe, David L Wood, Committee on Psychosocial Aspects of Child, Family Health, Committee on Early Childhood Adoption, Dependent Care et al.**, "The lifelong effects of early childhood adversity and toxic stress," *Pediatrics*, 2012, *129* (1), e232–e246. 2

**Siddique, Abu, Asad Islam, Tabassum Rahman, Tanvir Ahmed Mozumder, and Tanvir Shatil**, "Forced Displacement, Mental Health, and Child Development: Evidence from the Rohingya Refugees," *Working Paper*, 2022. 7

**Smith-Donald, Radiah, C Cybele Raver, Tiffany Hayes, and Breeze Richardson**, "Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research," *Early Childhood Research Quarterly*, 2007, *22* (2), 173–187. A11

**Unidad para las Víctimas**, "Registro Único de Víctimas," 2023. 14

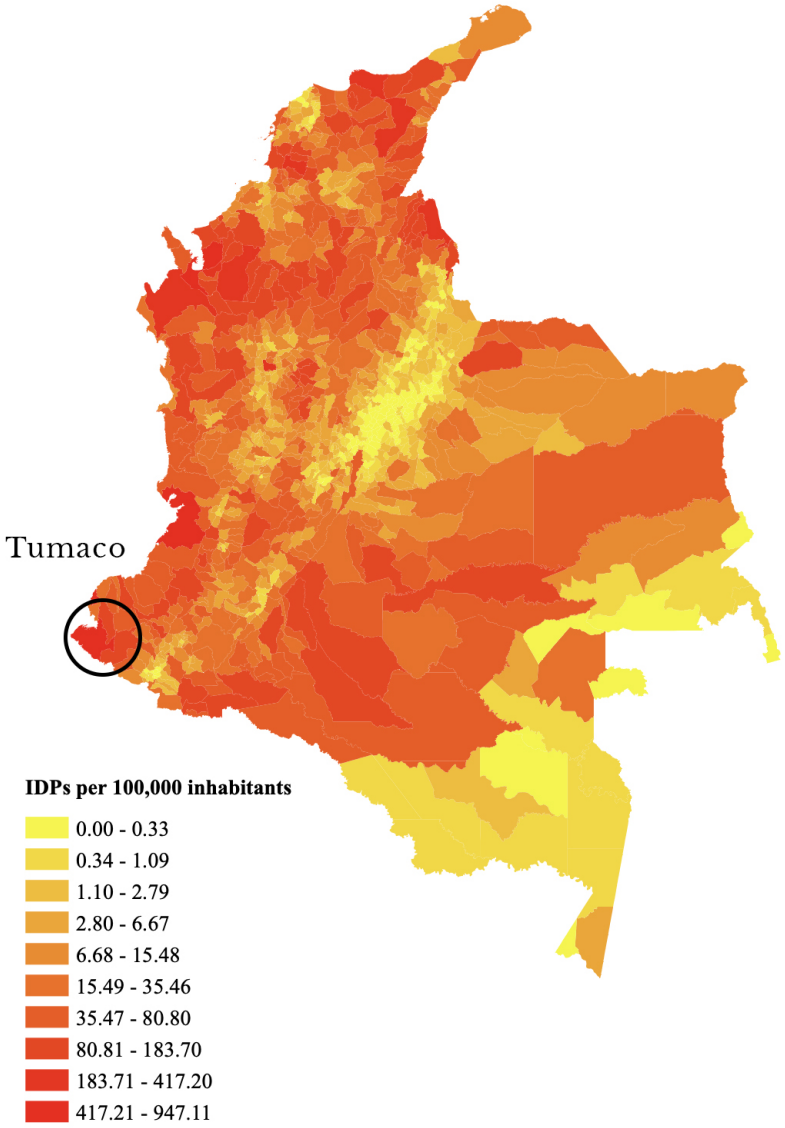
**United Nations High Commissioner for Refugees**, "Global Trends Forced Displacement in 2020," 2021. 14

**Wager, Stefan and Susan Athey**, "Estimation and inference of heterogeneous treatment effects using random forests," *Journal of the American Statistical Association*, 2018, *113* (523), 1228–1242. 40

**Zhou, Jin, James Heckman, Fuyao Wang, and Bei Liu**, "Early childhood learning patterns for a home visiting program in rural China," *Journal of Community Psychology*, 2022. 6

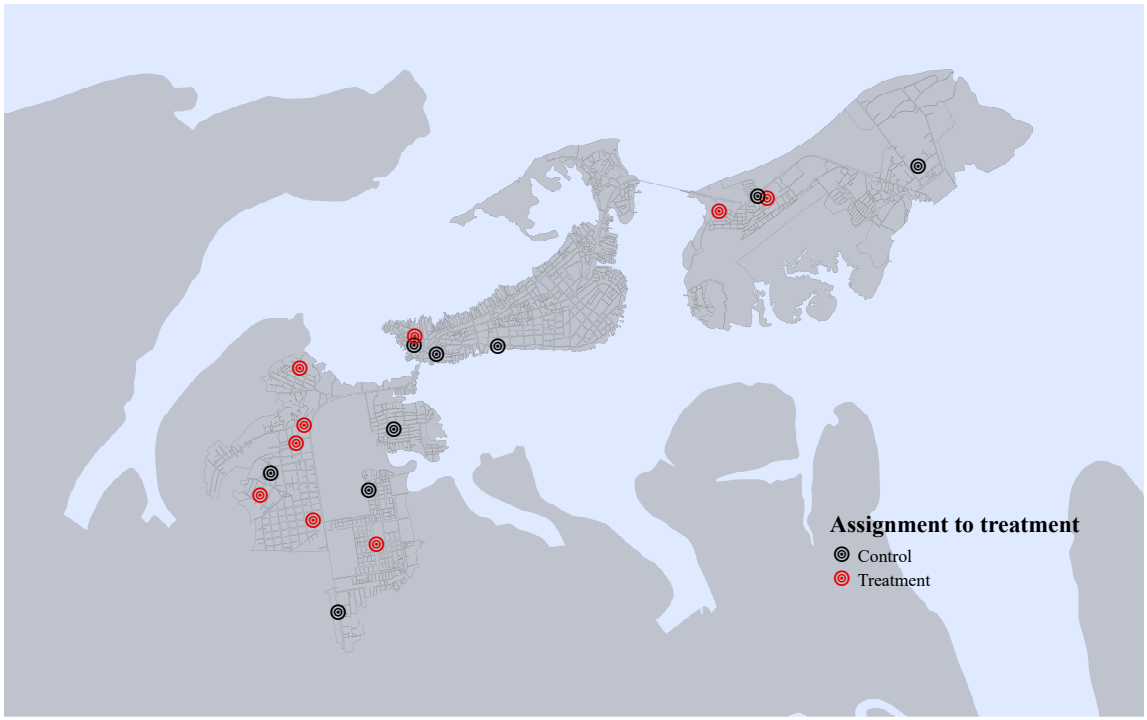
# Appendix

Figure A1: Intensity of Forced Displacement accross Municipalities in Colombia



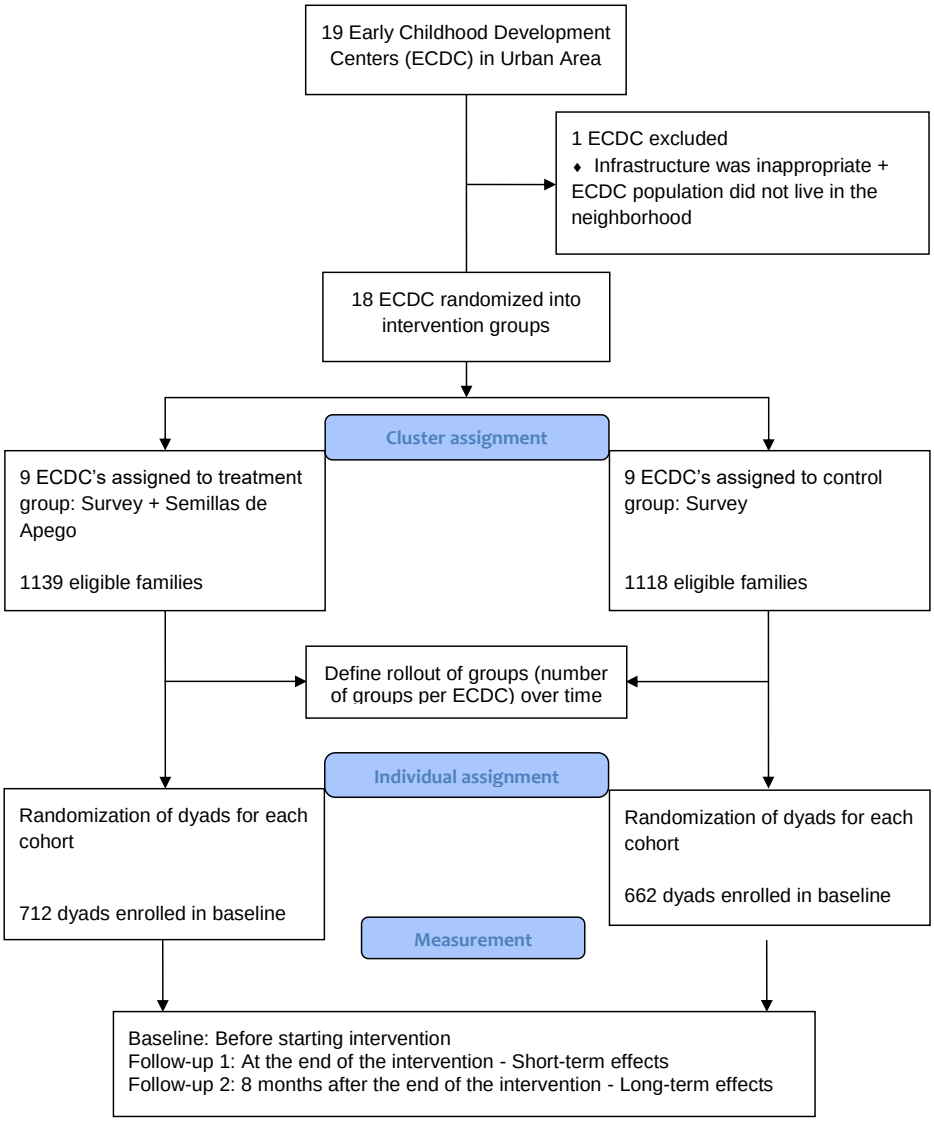

Notes: The map illustrates the intensity of forced displacement by municipality —number of IDPs per 100,000 inhabitants— based on data from the Registro Unico de Victimas, retrieved from https://www.unidadvictimas.gov.co/es/reportes on December 2021.

Figure A2: Childcare Centers in Tumaco




Notes: The map illustrates location of the childcare centers operated by Genesis Foundation, our partner in this study. These center serve approximately 1,600 children, corresponding to 80 percent of children served by public childcare centers in the urban area of Tumaco. Black icons denote centers assigned to the control arm of our experiment, while red icons denote centers assigned to the treatment arm.

Figure A3: Consort Diagram




Notes: The figure illustrates the consort diagram that summarizes the experimental design and the sample selection.

Figure A4: Participation in the program

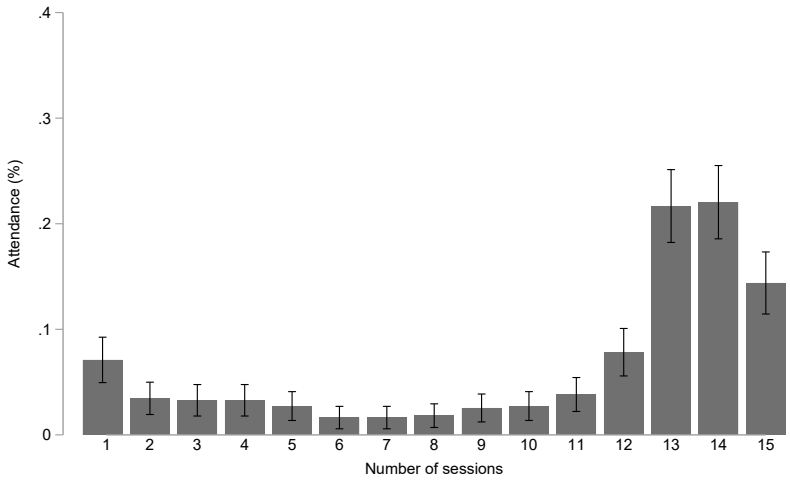


Figure A5: Participation in the program by cohorts

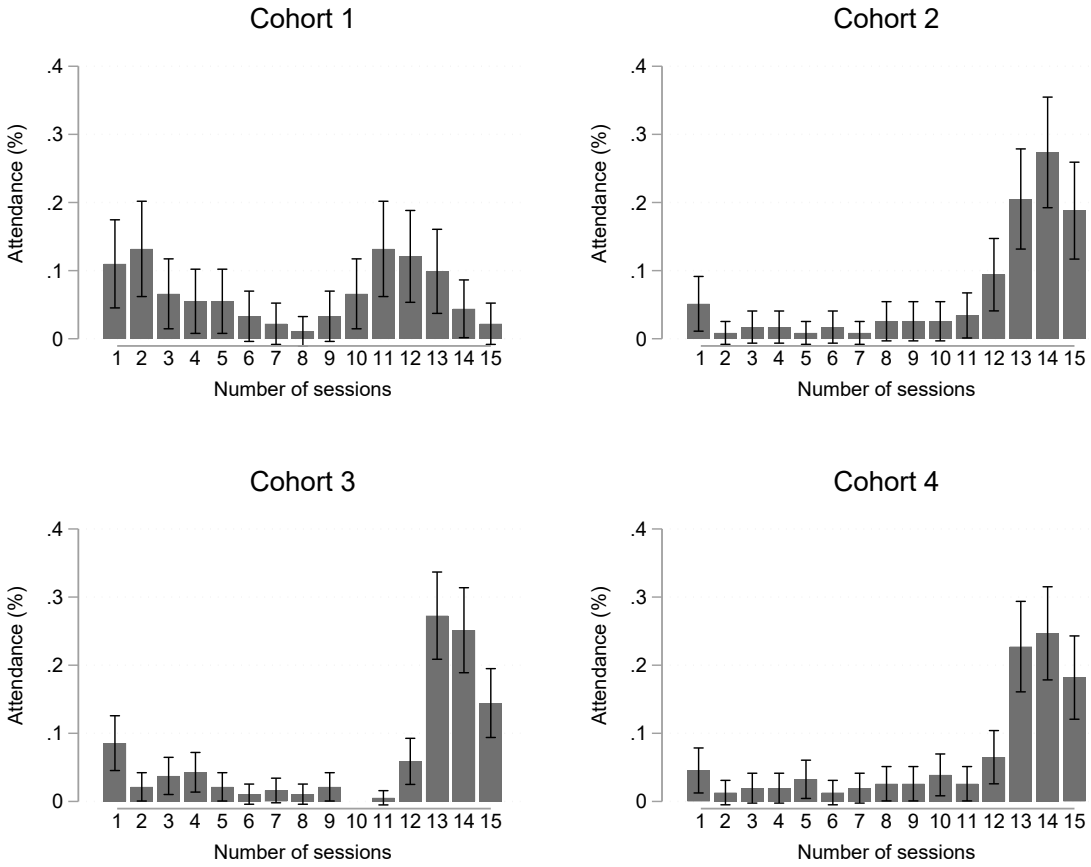



Figure A6: Timeline for data collection

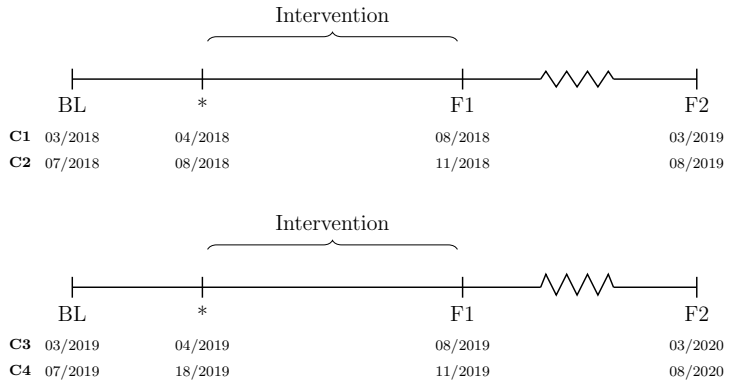


Figure A7: Timeline for data collection and Covid-19 interruption

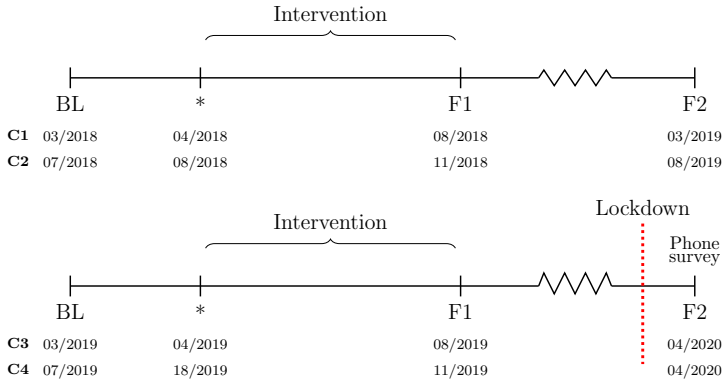



Table A1: Tumaco: Conflict and Socioeconomic Characteristics

| | Tumaco (1) | Colombia (2) |
|---|---|---|
| *Panel A. Poverty & demographics* | | |
| Multidimensional poverty index | 0.45 | 0.20 |
| Illiteracy rate | 0.14 | 0.08 |
| Low educational achievement | 0.47 | 0.41 |
| Inappropriate walls in dwellings | 0.45 | 0.10 |
| No health insurance | 0.17 | 0.10 |
| Formal employment | 0.08 | 0.27 |
| Afro-Colombian | 0.90 | 0.10 |
| *Panel B. Access to public services* | | |
| Unsafe drinking water | 0.44 | 0.11 |
| Sewerage | 0.08 | 0.92 |
| Aqueduct | 0.54 | 0.95 |
| Internet | 0.13 | 0.54 |
| Hospital beds (per 1,000) | 0.85 | 1.70 |
| Intensive care unit beds (per 1,000) | 0.01 | 0.11 |
| *Panel C. Violence* | | |
| Homicide rate (per 100,000) | 98.17 | 26.25 |
| Displacement - Expulsion rate per 100,000 | 2003 | 285 |
| Displacement - Reception rate per 100,000 | 943 | 238 |

Notes: Statistics from the National Administrative Department of Statistics and the National Planning Department.

Table A2: Average sessions attended: Unconditional and Conditional on attending to at least one session

| | Unconditional | Conditional |
|---|---|---|
| All | 8.31 | 10.79 |
| | (6.05) | (4.56) |
| Cohort 1 | 5.23 | 7.59 |
| | (5.24) | (4.68) |
| Cohort 2 | 9.23 | 11.91 |
| | (6.02) | (3.83) |
| Cohort 3 | 8.65 | 10.97 |
| | (6.13) | (4.70) |
| Cohort 4 | 9.27 | 11.62 |
| | (5.87) | (3.98) |

Notes: Standard deviations reported in parentheses.

Table A3: Selection at Baseline

| | Complete Sample | Cohorts1&2 | Cohorts3&4 |
|---|---|---|---|
| | (1) | (2) | (3) |
| Child age in months | 0.00 | -0.00 | 0.00 |
| | [0.002] | [0.003] | [0.002] |
| Female (=1) | 0.04 | 0.04 | 0.05 |
| | [0.027] | [0.042] | [0.036] |
| Child birth order | 0.00 | -0.01 | 0.02 |
| | [0.015] | [0.022] | [0.019] |
| Age | 0.00 | 0.00 | 0.00 |
| | [0.002] | [0.003] | [0.003] |
| Caregiver is the mother (=1) | 0.11* | 0.17* | 0.04 |
| | [0.050] | [0.066] | [0.072] |
| Household size | -0.01 | -0.02 | 0.00 |
| | [0.008] | [0.012] | [0.011] |
| Number of children under 5yrs | 0.06* | 0.09* | 0.03 |
| | [0.026] | [0.040] | [0.036] |
| Two-parent hh (=1) | 0.07* | 0.14** | 0.02 |
| | [0.031] | [0.047] | [0.041] |
| Highest years of education | 0.01 | 0.00 | 0.01 |
| | [0.005] | [0.007] | [0.006] |
| CCT beneficiary (=1) | 0.01 | -0.06 | 0.09* |
| | [0.029] | [0.045] | [0.038] |
| Head is employed (=1) | 0.00 | -0.02 | 0.04 |
| | [0.036] | [0.053] | [0.050] |
| Asset index | 0.02 | 0.05* | 0.00 |
| | [0.013] | [0.020] | [0.017] |
| Number of violent events | 0.02 | -0.00 | 0.03** |
| | [0.008] | [0.012] | [0.011] |
| IDP (=1) | 0.00 | 0.03 | -0.02 |
| | [0.031] | [0.049] | [0.040] |
| Caregiver Mental Health | -0.03 | -0.05 | -0.04 |
| | [0.021] | [0.032] | [0.026] |
| Child-Parent Relationships | -0.01 | -0.05 | 0.02 |
| | [0.019] | [0.028] | [0.025] |
| Child Mental Health | -0.00 | 0.05 | -0.02 |
| | [0.014] | [0.025] | [0.017] |
| Child Development | -0.02 | -0.01 | -0.04 |
| | [0.019] | [0.028] | [0.026] |
| Constant | 0.06 | 0.27 | -0.01 |
| | [0.143] | [0.224] | [0.183] |
| Observations | 1372 | 573 | 799 |
| R-squared | 0.027 | 0.063 | 0.049 |
| F-test | 2.222 | 2.613 | 2.647 |

Table A4: Childcare Characteristics

| | (1) Control | (2) Treatment | (3) Diff (2)-(3) |
|---|---|---|---|
| *Panel A. Randomization variables* | | | |
| ECDC size | 48.67 | 56.22 | -7.56 |
| | (22.417) | (56.835) | [20.365] |
| Children age 1 | 0.17 | 0.18 | -0.01 |
| | (0.092) | (0.076) | [0.040] |
| Children age 2 | 0.54 | 0.58 | -0.04 |
| | (0.125) | (0.100) | [0.053] |
| Children age 3 | 0.30 | 0.25 | 0.05 |
| | (0.125) | (0.086) | [0.051] |
| Male children | 0.50 | 0.48 | 0.02 |
| | (0.093) | (0.088) | [0.043] |
| SISBEN score | 15.56 | 15.73 | -0.17 |
| | (2.158) | (2.668) | [1.144] |
| Children's avg height | 89.05 | 89.45 | -0.40 |
| | (1.473) | (1.746) | [0.761] |
| Children's avg weight | 13.06 | 13.07 | -0.02 |
| | (0.315) | (0.468) | [0.188] |
| *Panel B. Other administrative information* | | | |
| Beneficiary is a victim of violence | 0.00 | 0.02 | -0.02 |
| | (0.006) | (0.047) | [0.016] |
| Beneficiary self-recognition of ethnic group | 0.45 | 0.74 | -0.30 |
| | (0.483) | (0.989) | [0.367] |
| Beneficiary belongs to Red de Unidos | 0.00 | 0.07 | -0.07 |
| | (0.004) | (0.189) | [0.063] |
| Mother is registered | 0.91 | 0.89 | 0.02 |
| | (0.163) | (0.187) | [0.083] |
| Father is registered | 0.52 | 0.56 | -0.05 |
| | (0.198) | (0.231) | [0.101] |
| Observations | 9 | 9 | 18 |

Notes:

Table A5: Intracluster Correlation Coefficients

| | ICC-All | ICC-Cohhort 1 | ICC-Cohort 2 | ICC-Cohort 3 | ICC-Cohort 4 |
|---|---|---|---|---|---|
| *Panel A. Child* | | | | | |
| Child Sex Female (=1) | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| Child birth order | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 |
| BITSEA score | 0.01 | 0.05 | 0.00 | 0.03 | 0.02 |
| TSCYC T-score | 0.01 | 0.03 | 0.02 | 0.03 | 0.03 |
| PSRA total score age stadarized | 0.02 | 0.01 | 0.00 | 0.00 | 0.09 |
| *Panel B. Caregiver* | | | | | |
| Female (=1) | 0.01 | 0.00 | 0.03 | 0.01 | 0.00 |
| Mother is caregiver (=1) | 0.00 | 0.00 | 0.02 | 0.00 | 0.02 |
| Literate (=1) | 0.02 | 0.04 | 0.01 | 0.09 | 0.00 |
| Highest years of education | 0.08 | 0.08 | 0.06 | 0.10 | 0.05 |
| Anxiety T-score | 0.01 | 0.07 | 0.00 | 0.05 | 0.04 |
| Depression T-score | 0.01 | 0.05 | 0.00 | 0.03 | 0.04 |
| Phobic anxiety T-score | 0.01 | 0.04 | 0.00 | 0.04 | 0.02 |
| Hostility T-score | 0.01 | 0.03 | 0.00 | 0.01 | 0.02 |
| Sensitivity T-score | 0.01 | 0.07 | 0.01 | 0.01 | 0.02 |
| Locus score | 0.02 | 0.03 | 0.04 | 0.04 | 0.03 |
| Severity Index T-score | 0.01 | 0.07 | 0.00 | 0.03 | 0.04 |
| PSI T-score | 0.01 | 0.07 | 0.01 | 0.00 | 0.04 |
| *Panel C. Household* | | | | | |
| Asset index | 0.02 | 0.07 | 0.03 | 0.03 | 0.00 |
| Household size | 0.01 | 0.00 | 0.00 | 0.03 | 0.05 |
| Two-parent hh (=1) | 0.01 | 0.00 | 0.05 | 0.02 | 0.01 |
| Family size | 0.01 | 0.00 | 0.09 | 0.00 | 0.00 |
| *Panel D. Violence* | | | | | |
| Number of violent events | 0.02 | 0.01 | 0.00 | 0.07 | 0.04 |
| Child exposure to violent events | 0.02 | 0.05 | 0.02 | 0.02 | 0.02 |
| Times hh has experienced force displacements | 0.03 | 0.03 | 0.04 | 0.02 | 0.03 |
| Victim of direct violence (=1) | 0.03 | 0.02 | 0.00 | 0.02 | 0.07 |
| Victimized child (=1) | 0.03 | 0.07 | 0.03 | 0.01 | 0.02 |

Table A6: Selection into each Cohort

| | Cohort | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Child Age in months | -0.005*** | 0.001 | 0.003** | 0.001 |
| | [0.001] | [0.001] | [0.001] | [0.001] |
| Female (=1) | 0.034 | 0.018 | -0.040 | -0.012 |
| | [0.022] | [0.022] | [0.025] | [0.024] |
| Child birth order | 0.004 | -0.002 | -0.018 | 0.016 |
| | [0.012] | [0.012] | [0.014] | [0.013] |
| Caregiver is the mother (=1) | -0.073 | 0.010 | 0.099* | -0.036 |
| | [0.051] | [0.053] | [0.059] | [0.057] |
| Female (=1) | -0.075 | -0.118* | 0.088 | 0.104 |
| | [0.069] | [0.071] | [0.080] | [0.077] |
| Age | -0.002 | -0.001 | 0.002 | 0.001 |
| | [0.002] | [0.002] | [0.002] | [0.002] |
| Years of education | -0.002 | 0.005 | -0.006 | 0.004 |
| | [0.006] | [0.006] | [0.007] | [0.006] |
| Household size | -0.002 | -0.006 | -0.000 | 0.008 |
| | [0.007] | [0.007] | [0.008] | [0.007] |
| Number of children under 5yrs | 0.045** | 0.009 | -0.056** | 0.002 |
| | [0.022] | [0.022] | [0.025] | [0.024] |
| Two-parent hh (=1) | -0.020 | 0.001 | 0.000 | 0.019 |
| | [0.025] | [0.026] | [0.029] | [0.028] |
| Highest years of education | 0.007 | -0.010 | 0.000 | 0.003 |
| | [0.006] | [0.006] | [0.007] | [0.007] |
| Asset index | -0.004 | -0.000 | 0.004 | 0.000 |
| | [0.011] | [0.011] | [0.012] | [0.012] |
| Hh income per capita | 0.000 | -0.000 | -0.000 | 0.000 |
| | [0.000] | [0.000] | [0.000] | [0.000] |
| CCT beneficiary (=1) | 0.091*** | 0.013 | 0.022 | -0.126*** |
| | [0.023] | [0.024] | [0.027] | [0.026] |
| Head is employed (=1) | -0.059** | 0.000 | 0.048 | 0.011 |
| | [0.029] | [0.030] | [0.034] | [0.033] |
| Head has formal job (=1) | 0.013 | -0.021 | -0.013 | 0.021 |
| | [0.028] | [0.029] | [0.033] | [0.032] |
| Victim of direct violence (=1) | -0.000 | -0.083** | 0.075* | 0.008 |
| | [0.036] | [0.037] | [0.041] | [0.040] |
| Number of violent events | -0.002 | 0.013 | 0.013 | -0.024*** |
| | [0.008] | [0.008] | [0.009] | [0.009] |
| IDP (=1) | 0.019 | 0.028 | -0.119*** | 0.071*** |
| | [0.024] | [0.025] | [0.028] | [0.027] |
| Caregiver Mental Health | -0.018 | -0.003 | 0.069*** | -0.049** |
| | [0.017] | [0.018] | [0.020] | [0.019] |
| Child-Parent Relationships | 0.064*** | -0.038** | -0.034* | 0.009 |
| | [0.015] | [0.016] | [0.018] | [0.017] |
| Child Development | -0.061*** | 0.083*** | 0.012 | -0.034** |
| | [0.015] | [0.016] | [0.018] | [0.017] |
| Child Mental Health | -0.004 | 0.018 | -0.010 | -0.004 |
| | [0.012] | [0.013] | [0.014] | [0.014] |
| Observations | 1372 | 1372 | 1372 | 1372 |
| Joint Significance F-test | 3.623 | 2.451 | 2.803 | 2.349 |

Notes: Columns (1) to (4) show the regression of whether being assigned to each cohort versus not being assigned to that cohort, e.g., cohort 1 is equal to 1 if the participant was assigned to cohort 1 and is equal to 0 if the participant was assigned to cohorts 2,3 or 4. Standard errors reported in brackets, * $\rho$ <0.1. ** $\rho$ <0.05. *** $\rho$ <0.01.

Table A7: Scales

| Dimension | Scale | Subscale | Source |
|---|---|---|---|
| Caregiver Mental Health | Questionnaire based on the Symptom Checklist 90-R (SCL-90) | Sensitivity<br>Depression<br>Phobic Anxiety<br>Anxiety<br>Hostility | Derogatis and Unger (2010) |
| | Locus of Control | | Rotter (1966) |
| | Inventory of Positive Psychological Attitudes (IPPA) | | Kass et al. (1991) |
| Child-Caregiver Relationships | Parental Stress Index (PSI - short form) | Difficult Child<br>Parent-Child Dysfunctional Interaction<br>Parental Distress | Haskett et al. (2006) |
| | Being a Mother (BaM) | Emotional closeness<br>Child Experience<br>Adult Experience | Matthey (2011) |
| | Child-Parent Relationships (CPR) | Conflicts<br>Closeness<br>Depedence | Driscoll and Pianta (2011) |
| | Interactions | Positive discipline<br>Negative discipline<br>Routines | Bernal et al. (2015) |
| Child Mental Health | Trauma Symptom Checklist for Young Children (TSCYC) | Anger<br>Anxiety<br>Sexual concern<br>Depression<br>Posttraumatic Stress<br>Dissociation | Briere et al. (2001) |
| Early Childhood Development | Preschool Self-regulation Assessment (PSRA) | | Smith-Donald et al. (2007) |
| | International Development Early Learning Assessment (IDELA) | Numeracy<br>Literacy<br>Socio-Emotional<br>Motor<br>Executive functions | Pisani et al. (2018) |
| | Brief Toddler Socioemotional Assessment (BITSEA) | Socio-Emotional / Behavioral Problems<br>Delays in Social-Emotional Competences | Briggs-Gowan and Carter (2007) |

Table A8: Psychometrics – Composite Indices

| | Cronbach's alpha | RMSEA |
|---|---|---|
| Caregiver's Mental Health | 0.952 | 0.06 |
| Child-Caergiver Relationships | 0.837 | 0.06 |
| Child Mental Health | 0.936 | 0.05 |
| Child Development (PSRA) | 0.711 | 0.06 |
| Child Development (Idela) - Only control | 0.809 | 0.05 |

Table A9: Psychometrics: Individual Scales

| | Cronbach's alpha | RMSEA |
|---|---|---|
| *Caregiver's Mental Health* | | |
| SCL | 0.952 | 0.06 |
| Locus | 0.647 | 0.10 |
| IPPA | 0.748 | 0.06 |
| *Child-Parent Relationships* | | |
| PSI | 0.885 | 0.07 |
| BAM | 0.632 | 0.07 |
| CPR | 0.731 | 0.08 |
| *Child-Parent Interactions* | | |
| Interactions | 0.569 | 0.08 |
| *Child Mental Health* | | |
| TSCYC | 0.936 | 0.05 |
| *Child Development* | | |
| BITSEA | 0.744 | 0.05 |
| PSRA | 0.692 | 0.06 |
| IDELA | 0.874 | 0.06 |

Table A10: Attrition

| | Cohort 1 | Cohort 2 | Cohort 3 | Cohort 4 | Total |
|---|---|---|---|---|---|
| *Panel A. Total* | | | | | |
| N | 277 | 296 | 429 | 374 | 1376 |
| Attrition rate at 1st followup | 0.03 | 0.04 | 0.05 | 0.05 | 0.04 |
| Attrition rate at 2nd followup | 0.06 | 0.11 | | | 0.09 |
| *Panel B. Treatment* | | | | | |
| N | 133 | 151 | 237 | 193 | 714 |
| Attrition rate at 1st followup | 0.03 | 0.03 | 0.03 | 0.07 | 0.04 |
| Attrition rate at 2nd followup | 0.08 | 0.09 | | | 0.08 |
| *Panel C. Control* | | | | | |
| N | 144 | 145 | 192 | 181 | 662 |
| Attrition rate at 1st followup | 0.03 | 0.05 | 0.06 | 0.03 | 0.04 |
| Attrition rate at 2nd followup | 0.05 | 0.13 | | | 0.09 |

Table A11: Balance: Treatment Vs Control for Non-Attritor Sample

| | (1) Control | (2) Treatment | (3) Diff (2)-(3) |
|---|---|---|---|
| *Panel A. Child* | | | |
| Child Age in months | 34.90 | 35.50 | -0.60 |
| | (8.205) | (8.470) | [0.459] |
| Female (=1) | 0.48 | 0.51 | -0.03 |
| | (0.500) | (0.500) | [0.028] |
| Child birth order | 1.88 | 1.93 | -0.05 |
| | (1.175) | (1.196) | [0.065] |
| *Panel B. Caregiver* | | | |
| Mother is caregiver (=1) | 0.85 | 0.88 | -0.03 |
| | (0.356) | (0.320) | [0.019] |
| Female (=1) | 0.94 | 0.97 | -0.04** |
| | (0.241) | (0.160) | [0.011] |
| Age | 28.93 | 29.29 | -0.36 |
| | (9.146) | (9.511) | [0.514] |
| Years of education | 11.90 | 11.88 | 0.02 |
| | (3.555) | (3.753) | [0.201] |
| *Panel C. Household* | | | |
| Household size | 4.99 | 5.03 | -0.04 |
| | (1.958) | (2.056) | [0.111] |
| Number of children under 5yrs | 1.26 | 1.32 | -0.06 |
| | (0.494) | (0.610) | [0.030] |
| Two-parent hh (=1) | 0.68 | 0.74 | -0.06* |
| | (0.467) | (0.440) | [0.025] |
| Highest years of education | 12.60 | 12.87 | -0.26 |
| | (3.323) | (3.269) | [0.182] |
| Asset index | -1.71 | -1.59 | -0.11 |
| | (1.184) | (1.200) | [0.066] |
| Hh income per capita | 246.90 | 235.56 | 11.35 |
| | (338.466) | (289.655) | [17.420] |
| CCT beneficiary (=1) | 0.42 | 0.44 | -0.02 |
| | (0.494) | (0.496) | [0.027] |
| Head is employed (=1) | 0.82 | 0.83 | -0.00 |
| | (0.382) | (0.379) | [0.021] |
| Head has formal job (=1) | 0.19 | 0.22 | -0.03 |
| | (0.394) | (0.418) | [0.022] |
| Days hh head works in a month | 23.51 | 23.92 | -0.41 |
| | (6.597) | (6.353) | [0.400] |
| *Panel D. Violence* | | | |
| Victim of direct violence (=1) | 0.80 | 0.84 | -0.04 |
| | (0.400) | (0.367) | [0.021] |
| Number of violent events | 2.25 | 2.49 | -0.24* |
| | (1.773) | (1.813) | [0.099] |
| IDP (=1) | 0.57 | 0.58 | -0.01 |
| | (0.496) | (0.494) | [0.027] |
| *Panel E. Indices* | | | |
| Caregiver Mental Health | 0.01 | -0.10 | 0.11* |
| | (0.755) | (0.821) | [0.043] |
| Child-Parent Relationships | -0.00 | -0.05 | 0.04 |
| | (0.819) | (0.852) | [0.046] |
| Child Mental Health | -0.01 | -0.10 | 0.08 |
| | (1.004) | (1.214) | [0.061] |
| Child Development | 0.00 | -0.04 | 0.05 |
| | (0.744) | (0.828) | [0.043] |
| Observations | 633 | 685 | 1318 |

Table A12: Attrition

| | F1 | | F2 | |
|---|---|---|---|---|
| Treatment | -0.003 | -0.004 | -0.007 | -0.008 |
| | [0.012] | [0.012] | [0.025] | [0.025] |
| Child age in months | 0.001 | 0.001 | 0.004* | 0.004 |
| | [0.001] | [0.001] | [0.002] | [0.002] |
| Female (=1) | 0.000 | 0.001 | 0.012 | 0.012 |
| | [0.011] | [0.011] | [0.024] | [0.023] |
| Child birth order | -0.006 | -0.006 | 0.021 | 0.021 |
| | [0.007] | [0.007] | [0.016] | [0.016] |
| Age | 0.000 | -0.000 | -0.002 | -0.002 |
| | [0.001] | [0.001] | [0.002] | [0.002] |
| Mother is caregiver (=1) | 0.023 | 0.022 | 0.006 | 0.002 |
| | [0.022] | [0.022] | [0.043] | [0.043] |
| Household size | -0.003 | -0.003 | -0.014* | -0.014* |
| | [0.003] | [0.003] | [0.007] | [0.007] |
| Number of children under 5yrs | 0.006 | 0.007 | -0.006 | -0.004 |
| | [0.011] | [0.011] | [0.020] | [0.020] |
| Two-parent hh (=1) | -0.032* | -0.032* | -0.025 | -0.027 |
| | [0.014] | [0.014] | [0.031] | [0.031] |
| Highest years of education | -0.005* | -0.005* | -0.010 | -0.009 |
| | [0.002] | [0.002] | [0.005] | [0.005] |
| CCT beneficiary (=1) | -0.006 | -0.005 | -0.048 | -0.045 |
| | [0.013] | [0.013] | [0.028] | [0.028] |
| Head is employed (=1) | 0.024 | 0.023 | 0.012 | 0.008 |
| | [0.014] | [0.014] | [0.034] | [0.034] |
| Asset index | 0.013* | 0.012* | 0.010 | 0.010 |
| | [0.006] | [0.006] | [0.012] | [0.012] |
| Number of violent events | -0.007 | -0.007 | 0.003 | 0.003 |
| | [0.004] | [0.004] | [0.008] | [0.008] |
| IDP (=1) | 0.013 | 0.013 | -0.038 | -0.038 |
| | [0.013] | [0.014] | [0.028] | [0.028] |
| Caregiver mental health index 3 raw scores std within wave | -0.006 | -0.006 | -0.008 | -0.009 |
| | [0.008] | [0.008] | [0.015] | [0.015] |
| Child-parent bond index 3 raw scores std within wave | -0.006 | -0.006 | 0.018 | 0.022 |
| | [0.008] | [0.008] | [0.013] | [0.014] |
| Early childhood development index 1 raw scores std within wave | 0.002 | 0.002 | 0.017 | 0.013 |
| | [0.006] | [0.006] | [0.013] | [0.013] |
| Cohort=2 | | 0.004 | | 0.042 |
| | | [0.016] | | [0.025] |
| Cohort=3 | | 0.011 | | |
| | | [0.016] | | |
| Cohort=4 | | 0.012 | | |
| | | [0.016] | | |
| Constant | 0.096 | 0.092 | 0.206 | 0.197 |
| | [0.059] | [0.059] | [0.141] | [0.141] |
| Observations | 1302 | 1302 | 527 | 527 |
| R-squared | 0.025 | 0.025 | 0.055 | 0.060 |
| F-test | 1.460 | 1.307 | 1.360 | 1.394 |
| Cohort Fixed Effect | No | Yes | No | Yes |

Table A13: Results at 1-month follow-up - Robustness

| | ITT | | | TOT |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A. Caregiver Mental Health* | | | | |
| Treatment | 0.016 | 0.022 | 0.009 | 0.015 |
| Cluster-Robust SE (CR2) | [0.037] | [0.035] | [0.029] | [0.047] |
| Q-value | (1.000) | (0.894) | (0.838) | (0.684) |
| Observations | 1317 | 1317 | 1317 | 1317 |
| *Panel B. Child-Parent Relationships* | | | | |
| Treatment | 0.016 | 0.018 | 0.045 | 0.075 |
| Cluster-Robust SE (CR2) | [0.065] | [0.064] | [0.056] | [0.091] |
| Q-value | (1.000) | (1.000) | (0.934) | (0.684) |
| Observations | 1318 | 1318 | 1318 | 1318 |
| *Panel C. Child-Parent Interactions* | | | | |
| Treatment | 0.208*** | 0.209*** | 0.209*** | 0.348*** |
| Cluster-Robust SE (CR2) | [0.046] | [0.038] | [0.035] | [0.058] |
| Q-value | (0.002) | (0.001) | (0.001) | (0.001) |
| Observations | 1318 | 1318 | 1318 | 1318 |
| *Panel D. Early Childhood Mental Health* | | | | |
| Treatment | -0.016 | -0.010 | 0.046 | 0.076 |
| Cluster-Robust SE (CR2) | [0.068] | [0.066] | [0.059] | [0.095] |
| Q-value | (1.000) | (1.000) | (0.934) | (0.684) |
| Observations | 1318 | 1318 | 1318 | 1318 |
| *Panel E. Early Childhood Development* | | | | |
| Treatment | 0.092 | 0.101 | 0.092 | 0.163 |
| Cluster-Robust SE (CR2) | [0.084] | [0.076] | [0.066] | [0.116] |
| Q-value | (1.000) | (0.665) | (0.269) | (0.482) |
| Observations | 950 | 950 | 950 | 950 |
| Lagged Dependent Variable | Yes | Yes | Yes | Yes |
| Cohort Fixed Effects | No | Yes | Yes | Yes |
| Controls | No | No | Yes | Yes |

Notes: Results of the lagged dependent variable model where the outcome indicated in each panel an measured at the 1-month followup is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of household controls in baseline. Controls for Panel A. include caregiver's age, asset index, two-parent household, number of violence events, whether the participant has been internally displaced, whether the household head had a formal job, and whether the participant had worked in the previous 7 days. Controls for Panel B. additionally include child's sex, child's birth order and caregiver's level of trauma. Lastly, controls for Panels C and D additionally include child's weight for height Z-score. Column 1 to 3 report the Intent-to-Treat effects, while Column 4 reports the Treatment-on-the-Treated effect for attending at least 6 sessions following a two-stage least squares model. Cluster robust standard errors are reported in brackets (CR2).; * $\rho$ <0.1. ** $\rho$ <0.05. *** $\rho$ <0.01.

Table A14: Results at 8-month follow-up - Robustness

| | ITT | | | TOT |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A. Caregiver Mental Health* | | | | |
| Treatment | 0.120* | 0.128** | 0.102* | 0.195* |
| Cluster-Robust SE (CR2) | [0.055] | [0.049] | [0.055] | [0.104] |
| Q-value | (0.070) | (0.046) | (0.078) | (0.060) |
| Observations | 522 | 522 | 522 | 522 |
| *Panel B. Child-Parent Relationships* | | | | |
| Treatment | 0.225* | 0.221** | 0.246** | 0.475*** |
| Cluster-Robust SE (CR2) | [0.080] | [0.082] | [0.075] | [0.149] |
| Q-value | (0.069) | (0.046) | (0.022) | (0.009) |
| Observations | 522 | 522 | 522 | 522 |
| *Panel C. Child-Parent Interactions* | | | | |
| Treatment | 0.139* | 0.143** | 0.110* | 0.213* |
| Cluster-Robust SE (CR2) | [0.078] | [0.073] | [0.065] | [0.124] |
| Q-value | (0.088) | (0.046) | (0.078) | (0.060) |
| Observations | 522 | 522 | 522 | 522 |
| *Panel D. Early Childhood Mental Health* | | | | |
| Treatment | 0.094 | 0.096* | 0.166* | 0.321* |
| Cluster-Robust SE (CR2) | [0.093] | [0.092] | [0.102] | [0.191] |
| Q-value | (0.134) | (0.093) | (0.078) | (0.060) |
| Observations | 522 | 522 | 522 | 522 |
| *Panel E. Early Childhood Development* | | | | |
| Treatment | 0.184* | 0.192** | 0.175** | 0.341** |
| Cluster-Robust SE (CR2) | [0.088] | [0.078] | [0.070] | [0.140] |
| Q-value | (0.074) | (0.046) | (0.046) | (0.050) |
| Observations | 505 | 505 | 505 | 505 |
| Lagged Dependent Variable | Yes | Yes | Yes | Yes |
| Cohort Fixed Effects | No | Yes | Yes | Yes |
| Controls | No | No | Yes | Yes |

Notes: Results of the lagged dependent variable model where the outcome indicated in each panel an measured at the 8-month followup is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of household controls in Baseline. Controls for Panel A. include caregiver's age, asset index, two-parent household, number of violence events, whether the participant has been internally displaced, whether the household head had a formal job, and whether the participant had worked in the previous 7 days. Controls for Panel B. additionally include child's sex, child's birth order and caregiver's level of trauma. Lastly, controls for Panels C and D additionally include child's weight for height Z-score. Column 1 to 3 report the Intent-to-Treat effects, while Column 4 reports the Treatment-on-the-Treated effect for attending at least 6 sessions following a two-stage least squares model. Cluster robust standard errors are reported in brackets (CR2).; * $\rho$ <0.1. ** $\rho$ <0.05. *** $\rho$ <0.01

Table A15: Heterogeneity: Forced Displacement

| | 1-month followup | | | 8-month followup | | |
|---|---|---|---|---|---|---|
| | Non-IDP | IDP | *pvalue* | Non-IDP | IDP | *pvalue* |
| Maternal Mental Health | 0.055 | 0.048 | 0.944 | 0.103 | 0.213* | 0.571 |
| | [0.082] | [0.072] | | [0.171] | [0.120] | |
| Observations | 564 | 753 | | 201 | 321 | |
| Child-Parent Relationships | 0.072 | 0.055 | 0.877 | 0.432** | 0.250** | 0.245 |
| | [0.090] | [0.068] | | [0.167] | [0.113] | |
| Observations | 564 | 754 | | 201 | 321 | |
| Child-Parent Interactions | 0.313*** | 0.195*** | 0.445 | 0.110 | 0.162 | 0.762 |
| | [0.093] | [0.071] | | [0.154] | [0.120] | |
| Observations | 564 | 753 | | 201 | 321 | |
| Child Mental Health | 0.190** | 0.056 | 0.146 | 0.137 | 0.220** | 0.650 |
| | [0.090] | [0.077] | | [0.159] | [0.108] | |
| Observations | 563 | 752 | | 199 | 319 | |
| Child Development | 0.121 | 0.293*** | 0.192 | 0.474*** | 0.214* | 0.188 |
| | [0.103] | [0.088] | | [0.165] | [0.115] | |
| Observations | 414 | 523 | | 195 | 309 | |

Table A16: Heterogeneity: Caregiver's Mental Health

| | 1-month followup | | | 8-month followup | | |
|---|---|---|---|---|---|---|
| | Not At-Risk | At-Risk | *pvalue* | Not At-Risk | At-Risk | *pvalue* |
| Maternal Mental Health | -0.325*** | 0.468*** | 0.000 | -0.167 | 0.677*** | 0.000 |
| | [0.083] | [0.095] | | [0.159] | [0.161] | |
| Observations | 658 | 659 | | 252 | 270 | |
| Child-Parent Relationships | 0.143 | 0.038 | 0.410 | 0.007 | 0.453*** | 0.015 |
| | [0.128] | [0.081] | | [0.210] | [0.138] | |
| Observations | 658 | 660 | | 252 | 270 | |
| Child-Parent Interactions | 0.084 | 0.296*** | 0.149 | -0.173 | 0.114 | 0.144 |
| | [0.123] | [0.090] | | [0.208] | [0.142] | |
| Observations | 658 | 659 | | 252 | 270 | |
| Child Mental Health | 0.002 | 0.115 | 0.430 | -0.122 | 0.483*** | 0.006 |
| | [0.109] | [0.103] | | [0.176] | [0.144] | |
| Observations | 657 | 658 | | 248 | 270 | |
| Child Development | 0.019 | 0.236** | 0.148 | 0.198 | 0.293** | 0.687 |
| | [0.147] | [0.104] | | [0.231] | [0.135] | |
| Observations | 483 | 454 | | 240 | 264 | |

Table A17: Heterogeneity: Socioeconomic Conditions: Asset Index Above or Below the Median

| | 1-month followup | | | 8-month followup | | |
|---|---|---|---|---|---|---|
| | Above median | Below median | *pvalue* | Above median | Below median | *pvalue* |
| Maternal Mental Health | 0.011 | -0.004 | 0.908 | 0.495** | -0.122 | 0.019 |
| | [0.112] | [0.124] | | [0.201] | [0.220] | |
| Observations | 659 | 658 | | 260 | 262 | |
| Child-Parent Relationships | 0.031 | 0.022 | 0.960 | 0.317 | 0.308 | 0.970 |
| | [0.118] | [0.119] | | [0.196] | [0.199] | |
| Observations | 659 | 659 | | 260 | 262 | |
| Child-Parent Interactions | 0.156 | 0.136 | 0.902 | 0.261 | 0.118 | 0.610 |
| | [0.121] | [0.125] | | [0.199] | [0.205] | |
| Observations | 659 | 658 | | 260 | 262 | |
| Child Mental Health | 0.114 | -0.073 | 0.380 | 0.236 | 0.074 | 0.485 |
| | [0.111] | [0.140] | | [0.169] | [0.219] | |
| Observations | 658 | 657 | | 259 | 259 | |
| Child Development | 0.211 | 0.085 | 0.423 | 0.123 | 0.393* | 0.210 |
| | [0.145] | [0.143] | | [0.209] | [0.209] | |
| Observations | 467 | 470 | | 256 | 248 | |

Table A18: Heterogeneity: Child age more or less than three years old

| | 1-month followup | | | 8-month followup | | |
|---|---|---|---|---|---|---|
| | Age ($\leq 3$) | Age ($> 3$) | *pvalue* | Age ($\leq 3$) | Age ($> 3$) | *pvalue* |
| Maternal Mental Health | 0.114* | -0.050 | 0.090 | 0.118 | 0.229 | 0.577 |
| | [0.064] | [0.077] | [0.112] | | [0.145] | |
| Observations | 775 | 542 | | 330 | 192 | |
| Child-Parent Relationships | 0.051 | 0.033 | 0.865 | 0.347*** | 0.099 | 0.174 |
| | [0.062] | [0.080] | | [0.104] | [0.139] | |
| Observations | 776 | 542 | | 330 | 192 | |
| Child-Parent Interactions | 0.237*** | 0.247*** | 0.911 | 0.146 | 0.086 | 0.484 |
| | [0.066] | [0.080] | | [0.104] | [0.148] | |
| Observations | 775 | 542 | | 330 | 192 | |
| Child Mental Health | 0.150** | 0.005 | 0.101 | 0.255** | -0.049 | 0.068 |
| | [0.066] | [0.086] | | [0.102] | [0.131] | |
| Observations | 774 | 541 | | 329 | 189 | |
| Child Development | 0.126 | 0.282*** | 0.305 | 0.309*** | 0.087 | 0.043 |
| | [0.078] | [0.095] | | [0.109] | [0.146] | |
| Observations | 561 | 376 | | 321 | 183 | |