Caregiver Mental Health and Early Childhood Development: Experimental Evidence from a Conflict-Affected Setting

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

April 24, 2024

Abstract

We report the results from a cluster-randomized control trial of a 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 had suffered conflict-related violence or had been forcefully displaced. Eight months after the program ended, we found positive intent-to-treat effects of 0.17 standard deviations (sd) on caregiver mental health; 0.31 and 0.15 sd on the quality and style of child-caregiver interactions, respectively; and 0.10 and 0.23 sd on early childhood mental health and development, respectively. Our findings speak to the importance of addressing caregivers' mental health as a binding *psychological constraint* for early childhood development in conflict-affected settings and other environments of adversity.

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

---

1 Introduction

A large body of interdisciplinary work emphasizes the importance of early childhood for brain development, skill formation, and physical and mental 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 quality of the relationship with an adult caregiver is one of the most important factors determining children's resilience and early development. However, exposure to traumatic experiences of conflict and forced displacement hurts caregivers' mental health (Charlson et al., 2019), 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 leads to the over-activation of different biological systems, including the body's stress-response system, a phenomenon frequently referred to as *toxic stress*. As we outline in more detail in the conceptual framework of Section 2, toxic stress affects brain architecture, neural threat-response, and emotional processing, hindering healthy early childhood development and compromising lifelong physical and mental health and socioeconomic trajectories (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 and other settings is limited.

To contribute towards closing these gaps, we report the results of the experimental evaluation of *Semillas de Apego*, a 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 and led by two community facilitators. As we describe in detail in Section 3, 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). 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 leabing them 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 its population and the highest figure worldwide (UNHCR, 2023; Unidad para las Víctimas, 2024).[1] Within Colombia, we conducted the impact evaluation in Tumaco, a municipality historically affected by conflict and forced displacement, and where the majority of its population has suffered from conflict-related violence and displacement, as we describe in Section 4. Baseline data highlights the pervasiveness of conflict in this context: 82 percent of the caregivers invited to participate in the study had experienced violence directly, 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 enrolled in public childcare centers that serve vulnerable and under-served families. We designed the experimental evaluation to overcome the challenges from the

---

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

ongoing conflict and the resource-constrained setting. Some of the elements in our design include the random assignment of the treatment at the childcare center level to overcome the 'invisible frontiers' set up by armed groups that prevent freedom of movement between neighborhoods; and 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 inference, and data quality. We discuss our experimental design and these challenges in Section 5.

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

We collected data at baseline and 1 and 8-month follow-ups on five dimensions following the program's theory of change: caregiver mental health, quality and style of the child-caregiver relationships, and children's mental health and early childhood development. For each dimension of interest, we first estimate the program's effects on a summary index following Kling et al. (2007). We also unpack the results by analyzing each scale and underlying dimension individually. In Section 6, we describe the scales we administered, their psychometric properties, and the data-collection protocols we implemented to overcome the challenges brought forth by the ongoing conflict. Unfortunately, the COVID-19 pandemic and the national lockdown implemented by the Colombian government prevented us from collecting the 8-month follow-ups for the last two cohorts of the program. For this reason, the analysis at the 8-month follow-up is based on 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 across the five dimensions of inter-

4

est. 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). 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) at the first follow-up. At the 8-month follow-up, we find sizeable and statistically significant improvements in the five core dimensions. Specifically, we find positive impacts of 0.17 standard deviations on the caregiver health index; 0.31 and 0.15 sd on the quality and style of child-caregiver interactions, respectively; and 0.10 and 0.23 sd on early childhood mental health development. These results suggest that the program's impacts persist and become larger over time, which is important given prior evidence of the fade-out of similar parenting interventions Bailey et al. (2020). We discuss these results in detail in Section 7 along with the results across the different scales and sub-dimensions.

We also find suggestive evidence of heterogeneity of impacts according to two dimensions of baseline vulnerability. The program has a stronger impact on caregivers with above-average symptoms of mental health problems at baseline. By contrast, the program had small impacts on caregivers who were food-insecure and had baseline levels of wealth below the sample median. 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).

Our work contributes to different strands of the literature. First, we build upon a large body of work in economics and 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 them 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. 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). And when maternal mental health is measured, it is often considered as a dimension of heterogeneity in impact. Together, this is indicative of how standard programming has ignored caregivers' mental health and its role for early childhood development. By contrast, our work pioneers by recognizing that in settings of conflict, forced displacement, and extreme adversities, caregivers' mental health is a binding *psychological constraint* that hinders their capacities to provide caring and nurturing environments for their children and that is detrimental for their development. Therefore, our results demonstrate the urgency and potential of addressing caregivers' mental health as a way to protect young children affected by conflict, forced displacement, and other adversities. 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 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 shocks or 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; Blattman et al., 2023; Bossuroy et al., 2022; Bhat 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 (Ibáñez and Moya, 2010a).[2] 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.[3] 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

---

[2]See Blattman and Miguel (2010) review on the causes and consequences of civil conflict (war).

[3]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)

trauma and poverty.

2   Conceptual Framework

In this section, we discuss the importance of caregivers' mental health as a binding *psychological constraint* in conflict-affected settings and other environments of extreme adversity. We use baseline data to describe the patterns of association between the severity of conflict-related violence and the caregiver's mental health, and the mediating role of the latter in explaining the effects of conflict on young children. Further, we discuss how the patterns in the data, even if descriptive, are consistent with the evidence in psychology and economics on the role of caregivers' mental health on early childhood development and motivate the specific approach of the program we evaluate in this paper.

In Figure 1 we plot the distributions of the caregiver mental health index at baseline according to the severity of pre-baseline conflict-related violence, defined as the number of violent events suffered. We stratify the data arbitrarily at the 75th percentile of the distribution for descriptive purposes but note that the patterns below are consistent at other points in the distribution. The figure illustrates the 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 extensively, including previous studies with internally displaced persons (IDPs) in Colombia (Moya, 2018) and survivors of mass violence (Mollica et al., 1998).

Figure 1: Conflict and Caregiver Mental Health
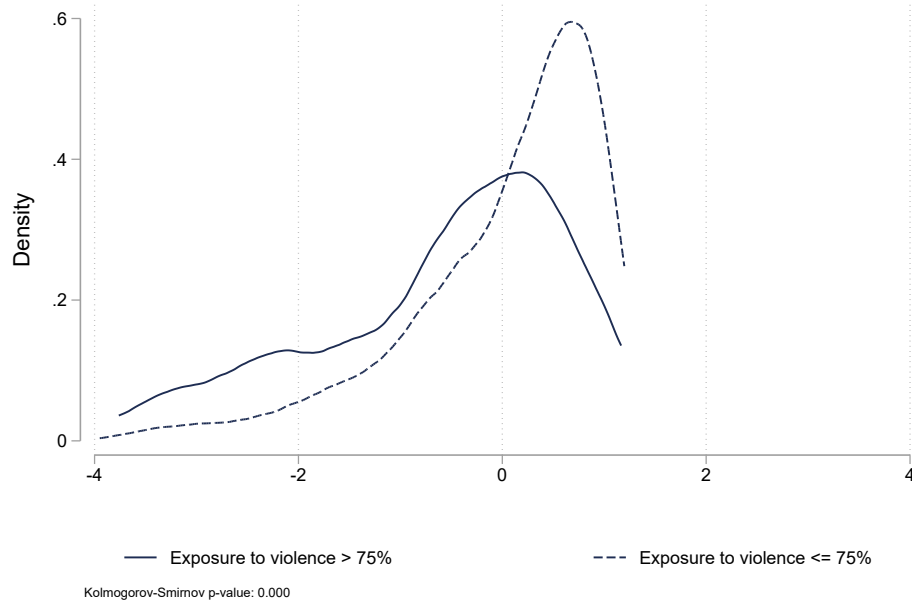
Notes: The figure plots the distribution of caregivers' mental health according to the severity of the exposure to violent events, defined as the number of violent events suffered). We stratify the data arbitrarily; above or below the 75th percentile of the distribution.

In Figure 2, we plot the distributions of four indices that capture the quality and style of the child-caregiver relationship and children's mental health and development. In these figures, we stratify the data according to whether the caregiver had at-risk symptoms of mental health problems at baseline.[4] The figures illustrate how caregivers' mental health is associated with worse quality and style of child-caregiver relationships and with worse children's mental health and development. Using this same data, Sánchez-Ariza et al. (2023) conduct a regression analysis, including a descriptive mediation analysis that indicates that caregivers' mental health explains between 55 and 78 percent of the direct effect of conflict on children's mental health.

The patterns in Figure 2 are consistent with existing evidence on how traumatic experiences affect early childhood development through their impact on caregivers' mental health. Mental health problems become a binding *psycholoigical constraint* as they affect

---

[4]As we discuss in Section 6, we consider at-risk symptoms for index scores that are at least one standard deviation below the baseline sample mean.

Figure 2: Caregiver Mental Health, Child-Caregiver Relationships, and Early Childhood Development
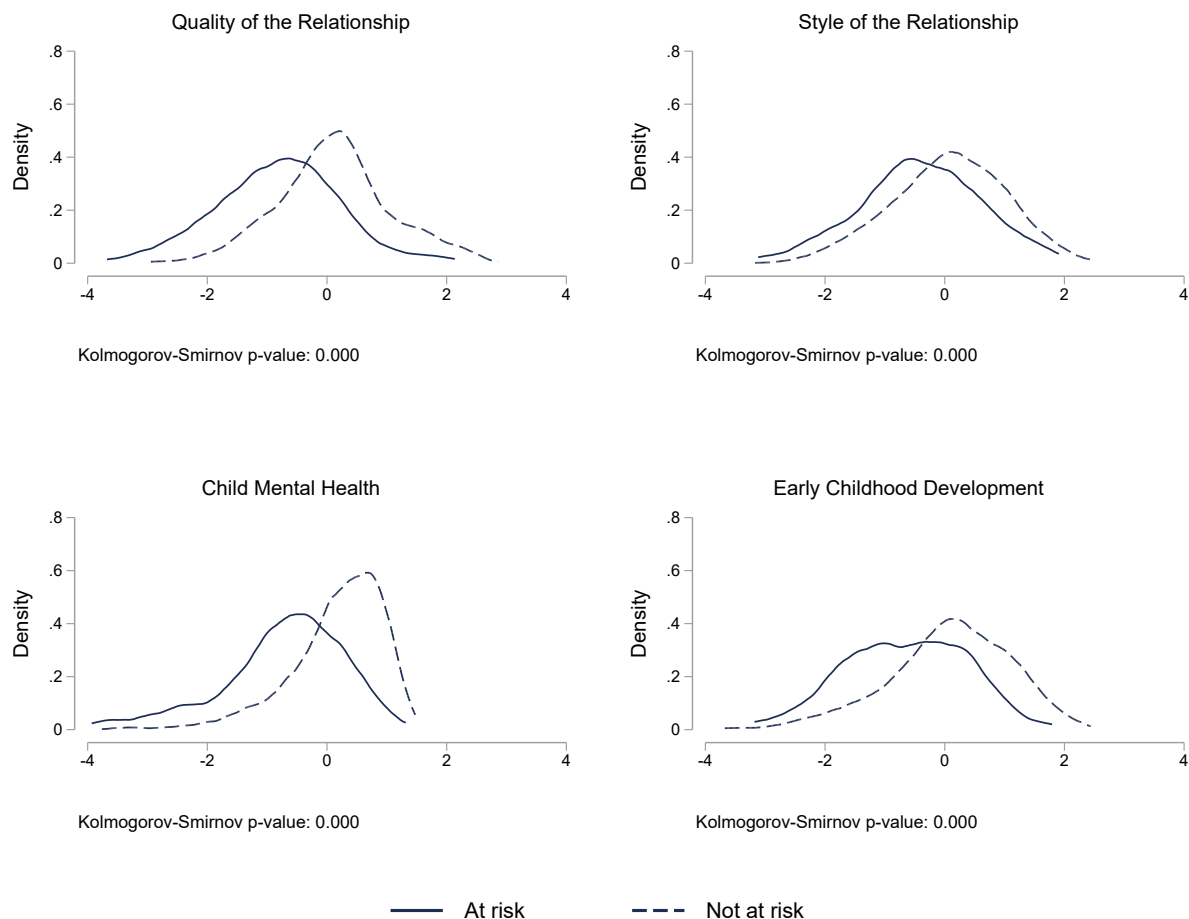
Notes: The figure plots the distribution of the quality and style of the child-caregiver relationship and of children's mental health and early childhood development according to the caregivers' symptoms of mental health problems (above or below at-risk levels).

caregivers' capacities to be emotionally available, hindering their ability to provide nurturing and developmentally appropriate caregiving, and leading, for example, to harsher parenting styles (Lieberman et al., 2006). The absence of these quality child-caregiver relationships, in combination with systematic early childhood adversities, affects young children by triggering *toxic stress*, the over-activation of different biological systems, which have severe consequences on children's brain architecture, neural threat response, and emotional processing (McLaughlin et al., 2019; Shonkoff et al., 2012; Nelson and Gabard-

[Durnam](), 2020). All of these effects have lasting effects on cognitive and socioemotional skill acquisition and physical and mental health and therefore set the seeds for the intergenerational transmission of mental health problems and poverty in conflict-affected settings.

3   Semillas de Apego

Semillas de Apego is a psychosocial community and group-based program for caregivers of children ages 0-5 in communities affected by armed conflict or forced displacement. Building upon the evidence outlined in the previous section, the program aims to restore caregivers' mental health as an outcome and as a pathway to promote healthy and nurturing child-caregiver relationships thereby protecting children's early development amid the adverse and traumatic experiences of conflict and armed displacement.

The program is implemented through 15 weekly sessions in groups of 15-20 caregivers (mothers, fathers, grandmothers, or other primary caregivers). Two community facilitators lead the sessions, which all follow a similar structure: (1) group discussion on progress over the prior week; (2) emotional regulation activity; (3) main activity: reflective practice based on art therapy components to ground the information; and (4) discussion on how to take the reflections and learnings into their regular lives and the child-caregiver relationship.

Semillas de Apego was designed building upon the Child-Parent Psychotherapy (CPP) (Lieberman and Van Horn, 2011) and Building Bridges (Reyes and Lieberman, 2010), two programs designed and implemented in the United States for underserved and vulnerable families, including migrant families enduring interpersonal violence or domestic abuse.[5]

---

[5]The CPP is a clinical therapeutic intervention implemented over an average of 30 sessions with the child-caregiver dyad and a therapist. Its efficacy has been validated in randomized trials with preschoolers exposed to domestic violence (Lieberman et al., 2005, 2006; Ippen et al., 2011); maltreated preschoolers (Toth et al., 2002); infants from families with a history of maltreatment (Cicchetti et al., 2006; Toth et al., 2015); anxiously attached infants of latinx immigrants (Lieberman et al., 1991); and toddlers with depressed caregivers (Cicchetti et al., 1999; Guild et al., 2017). The Building Bridges is a CPP group-based intervention that took CPP out of the clinical domain into a community setting but has not been evaluated yet.

Specifically, the program follows the theoretical framework of the CPP, which is based on attachment, cognitive-behavioral, and developmental theories, and trauma-informed models, and the group-based approach of Building Bridges. Further, the program's curriculum and implementation model were adapted to address the characteristics and needs of communities affected by ongoing conflict and forced displacement and facilitate its implementation in resourced-deprived settings.

The program is structured around a psychosocial intervention across four sequential themes (See Table 1). During the first 7 sessions, the program focuses on restoring caregivers' mental health, allowing recognition of how conflict, forced displacement, and other life experiences affect their mental health, and promoting emotional regulation and resilience. The following 6 sessions focus on understanding what young children need and how they are also affected by adversities, and fostering capacities to provide nurturing and developmentally appropriate care for their children. Finally, the last two sessions address social ecologies in which children grow to motivate caregivers to strengthen their support network and parenting teams.[6]

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.			

Semillas de Apego is delivered through a group-based model and a community-delivery implementation. These two components should not be understood as separate treatments

---

[6]Appendix B provides a detailed description of the program's conceptual framework, curriculum, and objectives

that influence the core outcomes independently, but rather as complements that facilitate the program's implementation and strengthen the psychosocial approach.

The community model is structured around the program's delivery by community facilitators, who do not have experience or training in psychosocial programs. This approach helps overcome the deficit in mental health services and the lack of trained professionals in conflict-affected settings, thereby contributing to the feasibility of the implementation. Furthermore, it contributes to lowering stigma towards mental health and enhancing community buy-in and caregivers' acceptability because facilitators can speak from their own experiences, which by and large are shared with the caregivers. To ensure the quality and fidelity of the implementation, community facilitators are trained through an experiential training model and supported throughout the programs' implementation through a reflexive supervision system.

In turn, the group-based model contributes to the program's feasibility and cost-effectiveness compared to an individual intervention. More importantly, it contributes to strengthening the objectives of the psychosocial approach. First, it tackles stigma by allowing caregivers to understand that other members of their community share mental health problems and are a normal reaction to abnormal and traumatic events. Second, it promotes peer learning by allowing caregivers to learn from the experiences, reflections, and ways other caregivers start incorporating new relationship models with their children. Finally, it helps restore the social connections that armed conflict and mental health problems often erode. In doing so, each group becomes a network that encourages adherence to the program sessions and provides a support system during and after the program's implementation.

4   Setting

Colombia has experienced a protracted armed conflict characterized by high levels of civilian victimization and forced displacement. By April 2024, the Colombian State has

legally recognized 9.7 million victims of the conflict, with 8.6 million IDPs (Unidad para las Víctimas, 2024). The latter represents 19 percent of all IDPs worldwide and 16 percent of the Colombian population (UNHCR, 2023). In addition, 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. Figure A1 in the Appendix 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 and in a region rich in natural resources provides suitable conditions for coca-leaf crops, drug trade, and illegal mining. As a result, armed groups have clashed and used violence towards civilians as a deliberate strategy for territorial control, consistent with conflict dynamics throughout the country (Grupo de Memoria Histórica, 2014). Conflict-related violence intensified after the largest guerrilla group (Fuerzas Armadas Revolucionarias de Colombia - FARC) signed a peace agreement in 2016 and other armed groups clashed for control of the areas left void.

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.[7] In addition, 45 percent of the population is below the national multidimensional poverty line, more than twice the national average.

---

[7]For comparison, Washington DC, one of the most violent cities in the United States, had a homicide rate of 40 in 2023.

5 Experimental Design

We now 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.

5.1 Implementation Strategy

We implemented the program in partnership with 18 public childcare centers in the urban area of Tumaco.[8] 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.[9] 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 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 this implementation strategy because the centers are considered safe havens and are protected by armed groups, even in the most violent neighborhoods. This allowed us to overcome the 'invisible borders' described above, enabled access to the different neighborhoods in Tumaco, and facilitated designing the sampling frame for the experiment using the centers' administrative data. This strategy also allowed us to build upon existing

---

[8]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.

[9]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)

relationships between the centers and the communities and encourage participation in the study in a setting where the conflict has eroded trust. Finally, this allowed us to host the sessions at the centers or in the nearby vicinity, which facilitated participation and contributed to ensuring the safety of participants, facilitators, and our research team.[10] However, by partnering with childcare centers, we missed out on caregivers of younger children (aged 0 to 2 years) who are not served by the centers but who may have benefited more from their caregivers' growth due to their developmental stage.

5.2 Experimental Design

Enrollment into the experiment was open to mothers or other primary caregivers of children aged 2 to 5 enrolled at the 18 childcare centers regardless of their prior exposure to conflict or mental health needs.[11] Over 4 sequential cohorts, 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. At the same time, caregivers in the treatment arm were invited to participate in the program.

We randomized access to the program following a two-stage cluster 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.[12] In each cohort, caregivers 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.[13] In treated

---

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

[11]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.

[12]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.

[13]When a caregiver had more than one child in the age range, we randomly selected one child to

centers, we then invited caregivers who had completed the survey to join the program. 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 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 considering the 'invisible borders', we would have endangered participants 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. Therefore, we chose to randomize treatment access at the child-care center level ($N_c = 18$) and considered the implications for inference and statistical power. For the latter, we pre-registered a minimum detectable effect (MDE) of 0.26 sd based on prior early childhood 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.[14] In section 7, we discuss how we estimate the clustered standard errors with a small number of clusters.

5.3 Descriptives

---

participate in the assessments.

[14]This is the estimate from our power calculations pre-registered before the baseline data were collected assuming 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 (See Appendix Table A2). Postregistry power calculations, after accounting for the observed ICCs and the variation explained by prognostic variables, yield a smaller MDE of 0.17 sd.

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	-0.08	0.00	-0.16	0.16**
	(1.065)	(1.000)	(1.118)	[0.057]
Quality of the Relationship	-0.04	0.00	-0.07	0.07
	(1.028)	(1.000)	(1.054)	[0.055]
Style of the Relationship	-0.02	0.00	-0.03	0.03
	(0.978)	(1.000)	(0.957)	[0.053]
Child Mental Health	-0.01	-0.00	-0.02	0.02
	(0.995)	(1.000)	(0.990)	[0.054]
Early Childhood Development	-0.05	0.00	-0.09	0.09
	(1.069)	(1.000)	(1.128)	[0.057]
Observations	1376	662	714	1376
Joint F-statistic	2.368			

Notes: Column 1 presents descriptive statistics for the entire sample. Columns 2 and 3 show the control and treatment arms, respectively, while column 4 shows the mean difference between these two groups. Standard deviations shown in parentheses and standard errors in brakets. * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$.

In Table 2, Column 1 we report descriptive statistics for all caregivers and index children in the study at baseline. On average, caregivers were 29 years of age, 96 percent were female, and 87 percent were the mothers of the index child. Children assessed in the study had 35 months of age at baseline on average and were the second child in the family. The data also highlights the socioeconomic vulnerability of the caregivers and their families. Caregivers had completed 11.85 years of education on average, only 21 percent had formal employment, and the average score in a multidimensional asset index was 1.64 standard deviations below the national mean.

As we discussed in the introduction, the data in the table also highlights the pervasiveness of the conflict. A majority of the caregivers in the study had directly experienced conflict-related violence or had been forcibly displaced. In addition, in Table AX, we find that 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.

5.4 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 A3 for unconditional and conditional attendance over the four cohorts and by cohort.[15] 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).[16] These patterns highlight high levels of participation and acceptability towards the program across the

---

[15]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 A4.

[16]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. 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)

four cohorts and signal 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.[17]

5.5   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 A5). We note a difference in the number of children enrolled at the centers (a 15 percent difference relative to the average size of control centers), which is explained by the random assignment of the largest childcare center to the treatment group. However, center size is not related to service quality as the same operator ran all centers, 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, Columns 2-4). Specifically, we achieved balance at baseline for 20 out of the 25 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 large 0.16 sd difference). Nevertheless, we fail to reject the null hypothesis that baseline characteristics jointly predict assignment into the treatment (F-test 2.37) as reported in Appendix Table A6. We will isolate

---

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

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 participants in the treatment centers who were assigned to the different cohorts were similar at baseline across the same dimensions described above. In Appendix Table A7, we show that baseline characteristics do not jointly predict assignment to a specific cohort (joint-significance F-tests range between 1.17 to 4.71).

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 observed that less than one percent of the caregivers who were reported in the center-level administrative data in 2018, 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, caregivers who were not served initially by the centers did not enroll 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. Yet, because the 'invisible borders' limit freedom of movement between different neighborhoods, we believe 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). Nevertheless, spillovers could have happened within treated centers across cohorts because of shared information between caregivers assigned to different cohorts or because of 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 have improved behavior and regulatory skills. Although we did not collect data on networks or interactions within treated centers, baseline comparisons between participants in treatment and

control centers in Cohorts 2-4 should provide a raw picture of such spillovers.

6 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.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.

6.1 Protocols

The 'invisible borders' also created significant risks for the data collection; not only to caregivers and their children but also to enumerators who would be at risk by visiting 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 reduce the number of caregivers and children we had to track outside of the centers during the 8-month follow up.[18] However, by limiting data collection at the centers, we could not collect 'objective'

---

[18]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.

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.

6.2 Outcomes

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 health, and early childhood development. For this purpose, we administered a battery of psychometric scales and assessments that had been validated or translated, and used previously in Colombia. Except for the observational evaluations of early childhood development, all data comes from scales answered by caregivers which could be affected by demand effects. Nevertheless, the scales inquire on multiple items or questions that point to an underlying construct, which helps control 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 not associated with the program to minimize demand effects. We describe each scale and assessment below and in Appendix Table A8, we provide more information on each instrument, including the main constructs and sub-constructs they capture.

Caregiver Mental Health: We assessed caregivers' mental health through a scale based on the Symptom Checklist 90-R (SCL-90R), which captures symptoms of different psy-

23

chopathologies: anxiety, phobic anxiety[19], depression, hostility,[20] and sensitivity.[21]

Quality of the Child-Caregiver Relationship: We administered two scales to measure the quality of the child-caregiver relationship: (1) The Parenting Stress Index - Short Form (PSI-SF) a scale that assesses the overall stress of the parenting relationship and captures subconstructs of parental distress, parent-child dysfunctional interactions, and difficult child; and (2) 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.[22] We use the data from these two scales to approximate different dimensions of the child-caregiver relationship including reciprocity and interplay. In addition, these two scales also capture the subjective experiences of caregivers, which the observational measures miss, and are important to understand in the context of the intervention.

Style of the Child-Caregiver Relationship: We administered a survey instrument based on the Colombian Longitudinal Survey (Bernal et al., 2015) measuring the type and frequency of caregivers interactions 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 quality and style of the child-caregiver relationships because the former conceptually captures the stress in the child-caregiver and attachment relationship. The latter measures the type 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

---

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

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

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

[22]We did not administer observational measures of the interactions between caregivers and their children, like the Strange 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.

dissociation. As discussed earlier, this data is reported by the caregiver.

Early Childhood Development: We evaluated children's development through developmentally appropriate assessments that rely on structured play-based tasks.[23] At baseline and the first post-intervention follow-up, we administered the Preschool Self-Regulation Assessment (PSRA), which measures self-regulation in emotional, attentional, and behavioral domains for children aged 30-40 months. At the 8-month follow-up, we administered the International Development Early Learning Assessment (IDELA), an assessment for children ages 3 to 6 that measures five 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.[24]

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 (Ibáñez and Moya, 2010b). 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 focuses on five composite indices, which we estimate by polling together all of the items from the scales and assessments in each dimension and estimating an index 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 all of the scales so that positive signs denote positive impacts (improvements) and

---

[23]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.

[24]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.

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 to the child-specific developmental trajectory. After analyzing the program's impact 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, we define within-sample risk thresholds as a score one standard deviation below the mean when they are not available.

6.3   Psychometrics

We assess the internal validity of the five composite indices by estimating the Cronbach Alphas and the Confirmatory Factor Analysis (CFA), which measure indices' internal consistency and whether their factor structure captures a latent construct, respectively. As a rule of thumb, a value greater than 0.70 in the Cronbach Alpha indicates an appropriate internal consistency. We found that four out of the five are internally valid at baseline with alphas ranging between 0.71 for the early childhood development index and 0.95 for the caregiver mental health index (see Appendix Table A9). The exception is the index of the style of child-caregiver relationships, which may be explained by the fact that this is not a scale in itself and measures different types of activities between caregivers and children that do not point in the same direction or are consistent among themselves. Nevertheless, 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.

In Appendix Table A10, 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 child-parent

relationship scale ($\alpha = 0.69$) with an alpha of 0.69 just below the rule-of-thumb threshold. In Column 2 we report the results from the CFA and observe a good fit in the RMSEA for all individual scales.

6.4   Covid-19 Disruption

We 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 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 have been able to collect the remaining data in person. We specified that if we could not go back to the field, our analysis 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 where we had expected larger effects because of the improvements in the dosage and quality of the implementation.

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 mental health consequences of the pandemic using this data.

6.5 Attrition

Finally, we analyze the rates of attrition across survey waves, the degree to which they differ across the treatment arms, and whether they are explained by baseline characteristics. We first found low rates of attrition: 96 and 91 percent of the caregivers who had completed the baseline assessment were surveyed in the 1 and 8-month follow-ups, respectively (see Appendix Table A11). 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.[25]

We find that attrition is not explained by differences in baseline characteristics. In Appendix Table A12 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 A13 we report the results from the joint-significance test and show that we fail to reject the null 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.

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

---

[25]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.

on the composite index for each of the five dimensions of interest of the program: caregiver mental health, quality and style of the child-caregiver relationships, child mental health, and early childhood development at each follow-up. For each dimension, we also 'unpack' the results on the individual scales and sub-scales. We estimate these effects at the intensive margin– the continuous z-scores provided by each scale— and at the extensive margin, the likelihood of at-risk symptoms. Because of the number of tests we conduct, we consider this analysis as exploratory and do not adjust for multiple hypothesis testing. To complement these results, and as a strategy to overcome the limitations in the statistical power from the small number of clusters, we conclude by reporting the results from an alternative empirical strategy where we exploit the phase-in design of the program and the random assignment of caregivers within the treatment centers to four sequential cohorts.

7.1 Empirical Strategy

We estimate the following lagged-dependent variable model for all indices, and for the scores for the individual scales and sub-scales:

$$Y_{ic}^f = \gamma_0 + \gamma_1 Treatment_c + \gamma_2 Y_{ic}^{bl} + \Gamma' X_{ic}^{bl} + \phi_k + \epsilon_i \tag{1}$$

where we regress the outcome Y_{ic}^f for individual i in childcare center c at the 1 or 8-month follow-up f, on the treatment assignment $Treatment_c$, the outcome at baseline Y_{ic}^f, a matrix X_{ic}^{bl} of baseline individual and household controls, and a set of cohort fixed effects ϕ_k. Following Imbens and Rubin (2015), we demean our covariates and fully interact them with the treatment indicator in our preferred specification.[26] Because we did not pre-register the set of covariates, we selected those where we found baseline imbalances (see Table 2) and standard demographic and socioeconomic controls.[27] For transparency,

---

[26]We also report the results of the more standard specification, controlling for the covariates separately.
[27]For caregiver outcomes, we control for their age and years of education, whether she has a formal job, whether she worked in the previous 7 days, whether she is part of a two-parent household, the household's asset index, years of education the number of conflict-related violent events suffered by the household, and whether the participant is an IDP. For outcomes on the child-caregiver relationship or the child, we

we report results with and without adjusted specifications. We cluster our standard errors at the childcare center level and follow Pustejovsky and Tipton (2018) to account for the small number of clusters. We also adjust our estimates for multiple hypothesis testing using Anderson's q-value (Anderson, 2008).

7.2 Preview of 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. The figure illustrates that at the 1-month follow-up, the program had sizeable and statistically significant impacts on the style of the child-caregiver relationship and early childhood development and moderate yet not statistically significant impacts on the quality of the relationships and child mental health. Considering the program's theory of change, the impact on early childhood development may come as a surprise because we observe a small and not statistically significant impact on caregivers' mental health. We discuss this in detail in the following subsection.

At the 8-month follow-up, we observe meaningful impacts in these two latter dimensions and stronger impacts on children's outcomes than at the first assessment. These results, in addition, are statistically different from zero even when the second follow-up analysis is based on a smaller sample because of the disruption from the COVID-19 pandemic, and confidence intervals are larger as a result. Taken together, the results suggest that the program's impacts build up over time, which is important given the evidence in education interventions on impacts that fade out in time (Bailey et al., 2020). Below, we analyze these results in detail.

---

in addition control for the child's sex, age (in days), birth order, weight for height index, and the caregiver mental health index at baseline (for which we documented baseline imbalance).

Figure 3: Impact at 1 and 8-month followups
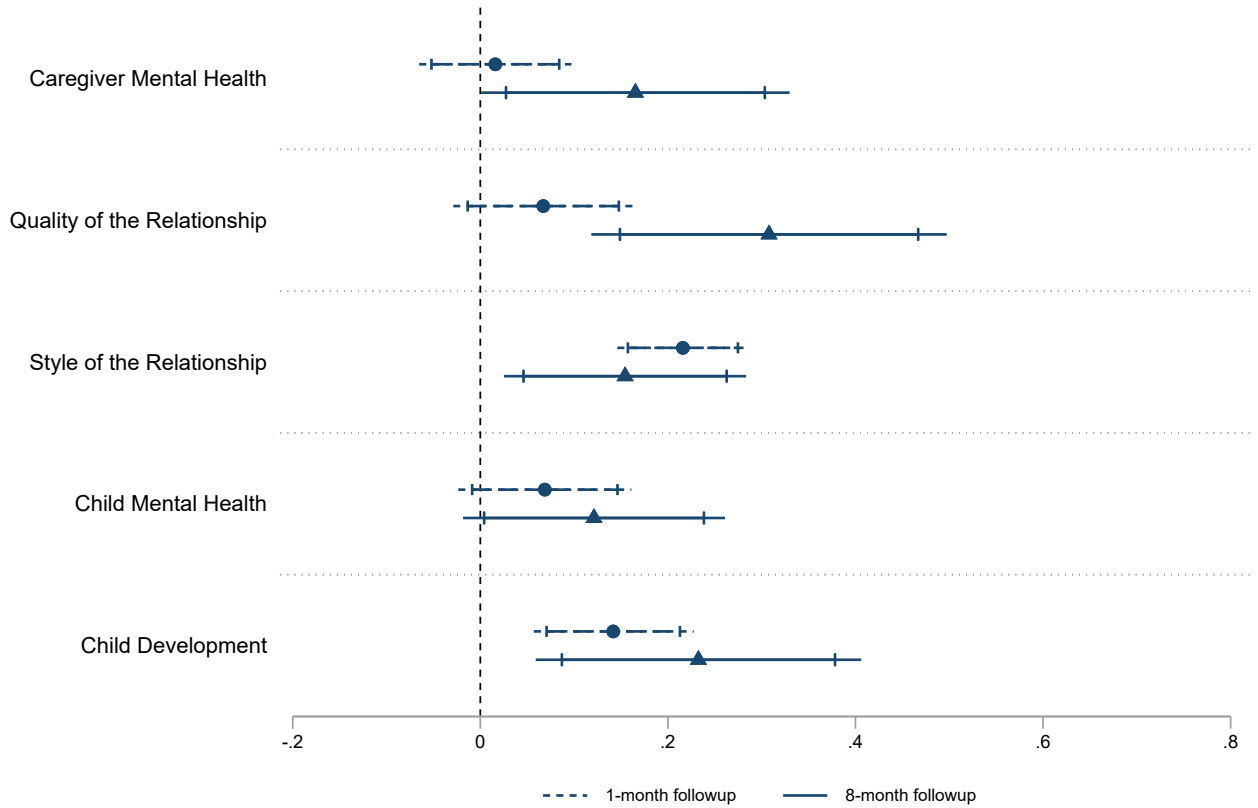
Notes: Graphical summary of the results of model 1 on the 1 and 8-month follow-ups on the five summary indices. 90% and 95% confidence intervals are reported in the bars and brackets.

7.3 Caregiver Mental Health

We report the results of the program's impact on the caregiver mental health index at the 1 and 8-month follow-ups in Table 3. This index can be understood as the global severity index of symptoms of mental health problems. In the table and in the ones that follow for the other indices, we report the ITT effect of model 1 when we only control for the lagged dependent variable (Column 1) and when we sequentially control for the cohort fixed effects and baseline individual and household covariates (Columns 2-3). In Column 4, we report the Average Treatment on the Treated (ATT) which we estimate instrumenting participation in 8 program sessions (the unconditional dosage) using the random assignment of the treatment at the childcare center. 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.

Table 3: Impact on Caregivers' Mental Health

	ITT			ATT
	(1)	(2)	(3)	(4)
Panel A. 1-month followup				
Caregiver Mental Health	0.020	0.026	0.016	0.027
	[0.046]	[0.045]	[0.041]	[0.063]
	(0.678)	(0.518)	(0.350)	(0.286)
Observations	1317	1317	1316	1316
Panel B. 8-month followup				
Caregiver Mental Health	0.192**	0.199**	0.165**	0.316**
	[0.077]	[0.077]	[0.084]	[0.153]
	(0.029)	(0.024)	(0.033)	(0.020)
Observations	522	522	522	522
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 (1) where the outcome indicated in each panel is measured at the 1 or 8-month followup is regressed on the treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of baseline covariates that are demeaned and fully interacted with the treatment status. Covariates include the caregiver's age, asset index, caregiver's years of education, 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. Columns 1-3 report the ITT, while Column 4 reports the ATT for caregivers who participated in at least 8 sessions (the unconditional dosage). CR2 cluster robust standard errors are reported in brackets, and Anderson's q-values for MHT are reported in parenthesis; * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$.

In Panel A of the table, we observe a small point estimate of the impact of the program at the 1-month follow-up on caregivers' mental health. This result is robust as we control for the cohort fixed effect and baseline covariates and for the ATT as well pointing to an imprecisely estimated null effect one month after the program ended. At face value, this result is at odds with the conceptual approach of the *Semillas de Apego* program and the positive impacts in the other dimensions that we previewed in Figure 3. If caregivers' mental health is the binding constraint for the promotion of nurturing child-caregiver relationships and of young children's mental health and development, it

is surprising that the program had no effect on caregivers' mental health but turned out to improve the child-parent relationship and children's early development. These results could suggest that the program improves these core outcomes but through mechanisms different than the improvement in caregiver mental health.

We argue, however, that these contradictory results can be explained by the program's 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 symptoms of mental health problems with a greater frequency because they become more aware relative to caregivers in the control arm, even if they are also experiencing improvements in their emotional regulation. Therefore, the point estimates of Table 3, Panel A may mask the improvements in mental health and emotional regulation with this greater awareness of symptoms of mental health problems. This hypothesis is consistent with the results of clinical studies of the CPP, which provided the conceptual basis for *Semillas de Apego* and, importantly, with our hypothesis registered in the pre-analysis plan. Further, similar dynamics have been observed in other psycho-education programs.

In Panel B of Table 3, we report the impacts on caregiver mental health at the second follow-up. We found that eight months after the program ended, the mental health index of caregivers assigned to the treatment arm denotes an improvement of 0.17 sd relative to those assigned to the control arm. This effect, in addition, is robust across the different specifications of model 1. Although the results are less precise because of the smaller sample, the estimated impact is also statistically significant by itself and after adjusting for multiple hypothesis testing (q-value = 0.033). Further, we observe a sizeable improvement of 0.32 sd for caregivers who participated in 8 or more sessions (q-value = 0.020).

To better understand these results, in Appendix Table A14 we unpack the mental health index and report the results on the individual subconstructs measured by the caregiver mental health scale. Consistent with the results above, we also find null impacts of

the program on the different psychopathologies at the first follow-up both at the intensive margin (Column 1) and the extensive margin (Column 3). We note that the point estimates for each subconstruct are negligible (smaller than 0.3 sd) pointing to imprecisely estimated null effects.

By contrast, we observe that at the 8-month follow-up, the program led to sizeable and statistically significant impacts between 0.152 and 0.181 sd in phobic anxiety, sensitivity, and hostility, with no apparent impacts on anxiety or depression. This result is important because the symptoms of phobic anxiety are more closely tied to the psychological consequences of conflict and forced displacement than depression and anxiety (Moya, 2018), while sensitivity and hostility are more closely connected with the dimensions of emotional regulation that the program targets (see Table 1). Furthermore, we find that at the 8-month follow-up, caregivers assigned to the treatment arm reported reductions of 5.9 percentage points (pp) in the likelihood of experiencing at-risk symptoms of phobic anxiety and sensitivity, and 3.5 pp in the likelihood of experiencing at-risk symptoms of hostility although the latter coefficient is only marginally statistically significant at conventional levels (p-value = 0.101). Again, the estimates at the 8-month follow-up are less precise than at the 1-month follow-up because of the smaller sample size used in this analysis but they still suggest a stronger impact consistent with the results above on the composite mental health index.

7.4   Child-Caregiver Relationships

We now turn our focus to the program's impacts on the quality and style of the child-caregiver relationship in Table 4. As we described before, we separate these two dimensions with the former capturing the stress and closeness in the relationship, while the latter is more descriptive, capturing the types of activities and discipline strategies caregivers use when interacting with the children.

In Panel A, we report the impacts on each composite index at the first follow-up.

Consistent with the results illustrated in Figure 3, we cannot reject the null hypothesis that the impacts on the quality of the relationship are different from zero. Across the different specifications, the point estimates in this dimension are small (< 0.74 sd), and not statistically significant. By contrast, we find a large ITT impact of 0.21 sd on the style of the child-caregiver relationship and a 0.359 sd impact for caregivers who attended more than 8 of the program's sessions.

Table 4: Impact on the Child-Caregiver Relationship

	ITT			ATT
	(1)	(2)	(3)	(4)
Panel A. 1-month followup				
Quality of the Relationship	0.070	0.074	0.067	0.111
	[0.065]	[0.066]	[0.049]	[0.075]
	(0.393)	(0.360)	(0.184)	(0.154)
Style of the Relationship	0.219***	0.221***	0.216***	0.359***
	[0.049]	[0.043]	[0.036]	[0.055]
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1317	1317	1316	1316
Panel B. 8-month followup				
Quality of the Relationship	0.260***	0.256***	0.308***	0.592***
	[0.095]	[0.096]	[0.097]	[0.174]
	(0.032)	(0.026)	(0.008)	(0.004)
Style of the Relationship	0.159**	0.165**	0.154**	0.299**
	[0.081]	[0.074]	[0.066]	[0.117]
	(0.037)	(0.026)	(0.027)	(0.015)
Observations	522	522	522	522
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 is measured at the 1 and 8-month follow-up is regressed on the treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include the caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, the caregiver's mental health, and child's age in days. Columns 1-3 report the ITT, while Column 4 reports the ATT for caregivers who participated in at least 8 sessions (the unconditional dosage). CR2 cluster robust standard errors are reported in brackets, and Anderson's q-values for MHT are reported in parenthesis; * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$.

In Panel B, we analyze the impacts on these two dimensions at the 8-month follow-up. Similar to what we observed for caregivers' mental health, we now identify a large and statistically significant ITT impact of 0.308 sd on the quality of the child-caregiver relationship, with an ATT of 0.592 sd. Furthermore, we still observe a positive and statistically significant impact on the style of the child-caregiver relationship, but we find a smaller impact than at the first follow-up with an ITT point estimate of 0.154 and an ATT estimate of 0.299.

In Tables Appendix A21 and A22 we unpack the results in these two dimensions across the differents scales and subscales. We find null impacts on the PSI and subconstructs at the intensive margin at the first follow-up, but find that the program led to reductions of 5.8 and 7.1 pp in the likelihood of at-risk symptons of parenting-stress (in the PSI global score) and in the caregiver's perception of the child being difficult, respectively. At this follow-up we also find improvements in the perceptions of emotional closeness between the caregiver and the chid, a subconstruct closely related to the quality of the attachment relationship, of 0.104 sd in the intensive margin and a reduction of 4.1 pp at the second follow-up.

We observe a similar picture at the 8-month follow-up with positive and statisticaly significant imapcts at the intensive margin on the PSI global score (0.215 sd), on the perceptions of the child being difficult (0.284 sd), and on the closeness of the child-caregiver relationship. Likewise, we find large reductions in the likelihood of experiencing at-risk symptoms of problems in the PSI global score (0.99 pp), difficult child (0.107 pp), and emotional closeness (0.039 pp).

7.5 Children's Mental Health and Development

Finally, we analyze the program's impacts on child menatl health and early childhood development in Table 5. We first find small and imprecisely estimated effects of 0.6 sd on childrens' mental health index but a sizeable and statistically significant impact of 0.142 sd

on the early childhood development index at the one-month follow-up. As we discussed before, the latter is measured with the BISTEA, focusing on children's socioemotional competences, and translates into a large 0.236 sd ATT.

Table 5: Impact on Children's Mental Health and Development

	ITT			ATT
	(1)	(2)	(3)	(4)
Panel A. 1-month followup				
Child Mental Health	0.026	0.031	0.069	0.115
	[0.056]	[0.052]	[0.047]	[0.074]
	(0.678)	(0.518)	(0.147)	(0.117)
				.
Observations	1317	1317	1316	1316
Child Development	0.131*	0.137**	0.142***	0.236***
	[0.074]	[0.067]	[0.043]	[0.070]
	(0.179)	(0.090)	(0.003)	(0.002)
Observations	1315	1315	1314	1314
Panel B. 8-month followup				
Child Mental Health	0.116*	0.122**	0.121*	0.232*
	[0.062]	[0.060]	[0.071]	[0.122]
	(0.029)	(0.024)	(0.037)	(0.024)
Observations	522	522	522	522
Child Development	0.262**	0.263**	0.233***	0.456***
	[0.109]	[0.107]	[0.089]	[0.162]
	(0.029)	(0.024)	(0.018)	(0.011)
Observations	504	504	504	504
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 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 controls in the baseline. Controls include the caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. The child's age in days is excluded as a control for Child Development since the index is already standardized with respect to the child's age. Columns 1 and 3 report the ITT, while Column 4 reports the ATT effect for attending the average number of sessions following a two-stage least squares model. Cluster robust standard errors are reported in brackets (CR2); * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$.

Consistent with the majority of the results in the previous dimensions and with the programs theory of change, at the second follow-up we find meaningful impacts on the child mental health and early childhood development, which now combines the caregiver-reported BITSEA and the IDELA child assessment. In particular, we identify an impact of 0.105 sd on children's mental health, which is not statistically significant when only adjusting for the CR2 standard error but does appear different from zero once we take into account the multiple hypothesis adjustment, which happens because the proportion of null hypothesis is lower than the expected for the corresponding p-value. The ATT in this case corresponds to a large impact of 0.201.

Likewise, we find large and robust impact on early childhood development of a magnitude of 0.233 sd for the ITT and 0.456 sd for the ATT. The former captures a sizeable and meaningful effect, simlar of those of standard parenting interventions in the region despite the more profound vulnerabilities in this context, and the perhaps larger risks caregivers and children face including the on-going conflict. In addition, the results in this domain, similar to those of other dimensions indicate that the program had impacts that grew and persisted over time, consistent with the conceptual framework and theory of change. Further work should 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.

In Tables Appendix A23 and A24 we analyze the impacts across the different psychopathologies of children's mental health and dimensions of early childhood development respectively. In Table Appendix A23 we identify a positive impact on anxiety at both the intensive (0.101 sd) and extensive (-0.36 pp) margins. More importantly, at the second follow-up we observe sizeable and significant impacts especially at the extensive margin, lowering the likelihood of at-risk symptoms of anger (-0.60 pp), anxiety (-0.095 ppp),

depression (-0.083), sex concern (-0.065 pp), and post-traumatic stress (0.085 pp).

In Table Appendix A24 we identify positive and sizeable impacts on the BITSEA global score (0.144 sd) and socioemotional learning (0.120 sd) at the fist assessment post-intervention. These two effects largely explain the large impact we had observed in early childhood development in Table 5. In this same follow-up, we also note a large 22.7 pp reduction in the likelihood of at-risk socioemotional problems. While these intensive-margin effects seem to fade out in the second follow-up, we nevertheless identify sizeable impacts at the extensive margin for the BITSEA global score and each individual component – reduction of at-risk lielihood of 0.089 pp in the BITSEA global score, and 0.064 and 0.66 in the socioemotional and behavioral problems subscales, respectively. By contrast, we obtain null or noisy point estimators for the IDELA assessment, both at the extensive and intensive margin. For the IDELA, we only note a large improvement of 0.295 sd in children's executive functions. Further, we do not find posirtive impacts at the intensive or extensive margin for the IDELA's socioemotional component, a results that is consistent with the estimated impacts from the BITSEA which, at the second follow-up, emerged for the global score and bhevaioral problems dimension but not for socioemotional skills.

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 A17-A20). 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
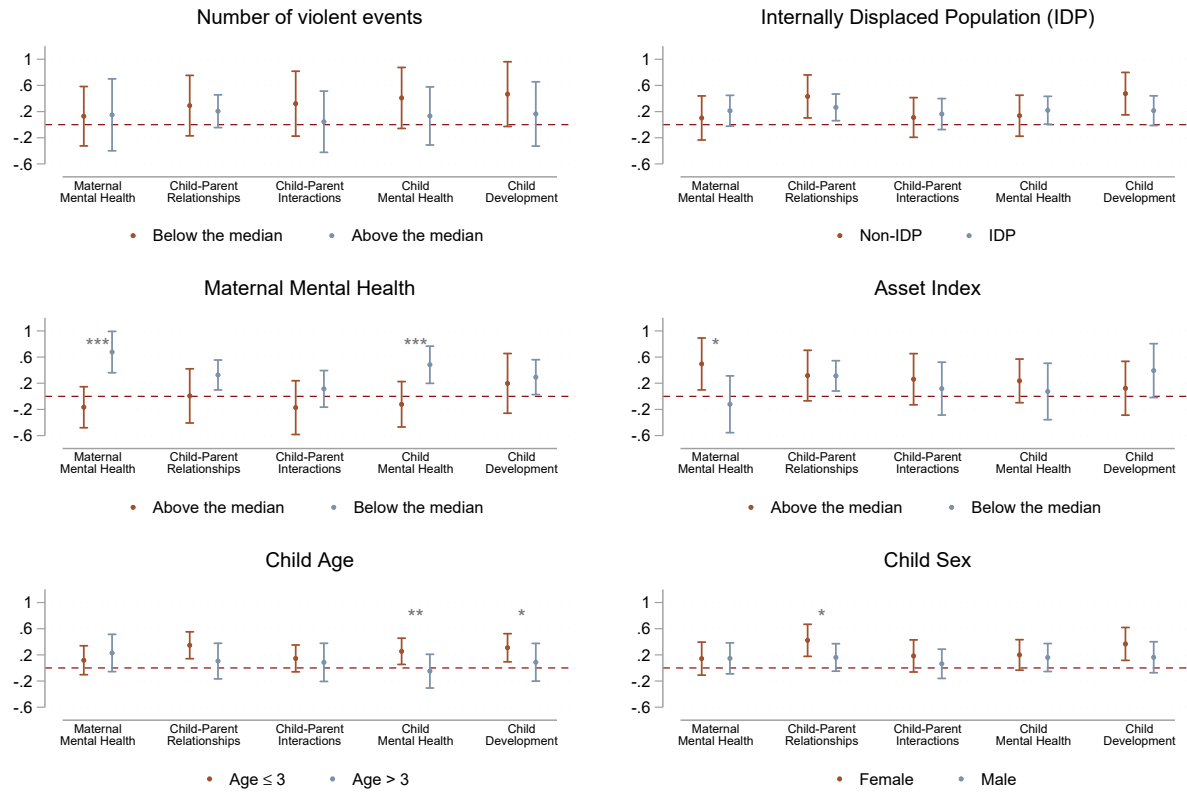
Statistical difference: * p < .01, ** p < .05, *** p < .1

7.7 Robustness: Within-treatment variation.

Finally, we exploit 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 6. 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 follow-up.

Table 6: Results: Within-Treatment (ITT)

	ITT	
	1-month	8-month
Panel A. Caregiver Mental Health		
Treatment	0.222***	0.268***
	[0.079]	[0.086]
	(0.011)	(0.002)
Observations	702	689
Panel B. Quality of the Relationships		
Treatment	0.075	0.378***
	[0.071]	[0.079]
	(0.196)	(0.001)
Observations	703	689
Panel B. Style of the Relationsips		
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. 30

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. 7

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. 17

_ , **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

Bailey, Drew H, Greg J Duncan, Flávio Cunha, Barbara R Foorman, and David S Yeager, "Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions," *Psychological Science in the Public Interest*, 2020, *21* (2), 55–97. 5, 30

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. 7

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, A15

_ , **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

Bhat, Bhargav, Jonathan De Quidt, Johannes Haushofer, Vikram H Patel, Gautam Rao, Frank Schilbach, and Pierre-Luc P Vautrey, "The long-run effects of psychotherapy on depression, beliefs, and economic outcomes," Technical Report, National Bureau of Economic Research 2022. 7

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. 7

_ , **Sebastian Chaskel, Julian C Jamison, and Margaret Sheridan**, "Cognitive Behavioral Therapy Reduces Crime and Violence over Ten Years: Experimental Evidence," *American Economic Review: Insights*, 2023, *5* (4), 527–545. 7

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. 7

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 multisite study," *Child Abuse & Neglect*, 2001, *25*, 1001–1014. A15

Briggs-Gowan, Margaret J and Alice S Carter, "Brief infant toddler social emotional assessment," 2007. A15

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

Charlson, Fiona, Mark van Ommeren, Abraham Flaxman, Joseph Cornett, Harvey Whiteford, and Shekhar Saxena, "New WHO prevalence estimates of

mental disorders in conflict settings: a systematic review and meta-analysis," *The Lancet*, 2019, *394* (10194), 240–248. 2

Cicchetti, Dante, Fred A Rogosch, and Sheree L Toth, "Fostering secure attachment in infants in maltreating families through preventive interventions," *Development and psychopathology*, 2006, *18* (3), 623–649. 11

_ , Sheree L Toth, and Fred A Rogosch, "The efficacy of toddler-parent psychotherapy to increase attachment security in offspring of depressed mothers," *Attachment & Human Development*, 1999, *1* (1), 34–66. 11

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. A15

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). A15

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. 5

Evans, David K, Pamela Jakiela, and Heather A Knauer, "The impact of early childhood interventions on mothers," *Science*, 2021, *372* (6544), 794–796. 6

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

Guild, Danielle J, Sheree L Toth, Elizabeth D Handley, Fred A Rogosch, and Dante Cicchetti, "Attachment security mediates the longitudinal association between child–parent psychotherapy and peer relations for toddlers of depressed mothers," *Development and psychopathology*, 2017, *29* (2), 587–600. 11

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. A15

Haushofer, Johannes and Ernst Fehr, "On the psychology of poverty," *science*, 2014, *344* (6186), 862–867. 6

Ibáñez, Ana María and Andrés Moya, "Do conflicts create poverty traps? Asset losses and recovery for displaced households in Colombia," in "The economics of crime: Lessons for and from Latin America," University of Chicago Press, 2010, pp. 137–172. 7

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

Imbens, Guido W and Donald B Rubin, *Causal inference in statistics, social, and biomedical sciences*, Cambridge university press, 2015. 29

Ippen, Chandra Ghosh, William W Harris, Patricia Van Horn, and Alicia F Lieberman, "Traumatic and stressful events in early childhood: Can treatment help those at highest risk?," *Child abuse & neglect*, 2011, *35* (7), 504–513. 11

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. 6

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119. 4, 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, 11

_ , Chandra Ghosh Ippen, and Patricia Van Horn, "Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial," *Journal of the American Academy of Child & Adolescent Psychiatry*, 2006, *45* (8), 913–918. 10, 11

_ , Donna R Weston, and Jeree H Pawl, "Preventive intervention and outcome with anxiously attached dyads," *Child development*, 1991, *62* (1), 199–209. 11

_ , **Patricia Van Horn, and Chandra Ghosh Ippen**, "Toward evidence-based treatment: Child-parent psychotherapy with preschoolers exposed to marital violence," *Journal of the American Academy of Child & Adolescent Psychiatry*, 2005, *44* (12), 1241–1248. 11

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. 5

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, 10

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. 8

Moya, Andrés, "Violence, psychological trauma, and risk attitudes: Evidence from victims of violence in Colombia," *Journal of Development Economics*, 2018, *131*, 15–27. 6, 8, 19, 24, 34

_ , **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. 7

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, 10

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. A15

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. 30

Reyes, Vilma and Alicia Lieberman, "Building Bridges: Child-Parent Psychotherapy Group Model," 2010. 11

Ridley, Matthew, Gautam Rao, Frank Schilbach, and Vikram Patel, "Poverty, depression, and anxiety: Causal evidence and mechanisms," *Science*, 2020, *370* (6522), eaay0214. 6

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. 20

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. 9

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, 10

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. 6

Toth, Sheree L, Angeline Maughan, Jody Todd Manly, Mary Spagnola, and Dante Cicchetti, "The relative efficacy of two interventions in altering maltreated preschool children's representational models: Implications for attachment theory," *Development and psychopathology*, 2002, *14* (4), 877–908. 11

_ , Melissa L Sturge-Apple, Fred A Rogosch, and Dante Cicchetti, "Mechanisms of change: Testing how preventative interventions impact psychological and physiological stress functioning in mothers in neglectful families," *Development and psychopathology*, 2015, *27* (4pt2), 1661–1674. 11

UNHCR, "Global Trends: Forced Displacement in 2022," 2023. 3, 14

Unidad para las Víctimas, "Registro Único de Víctimas," 2024. 3, 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. 39

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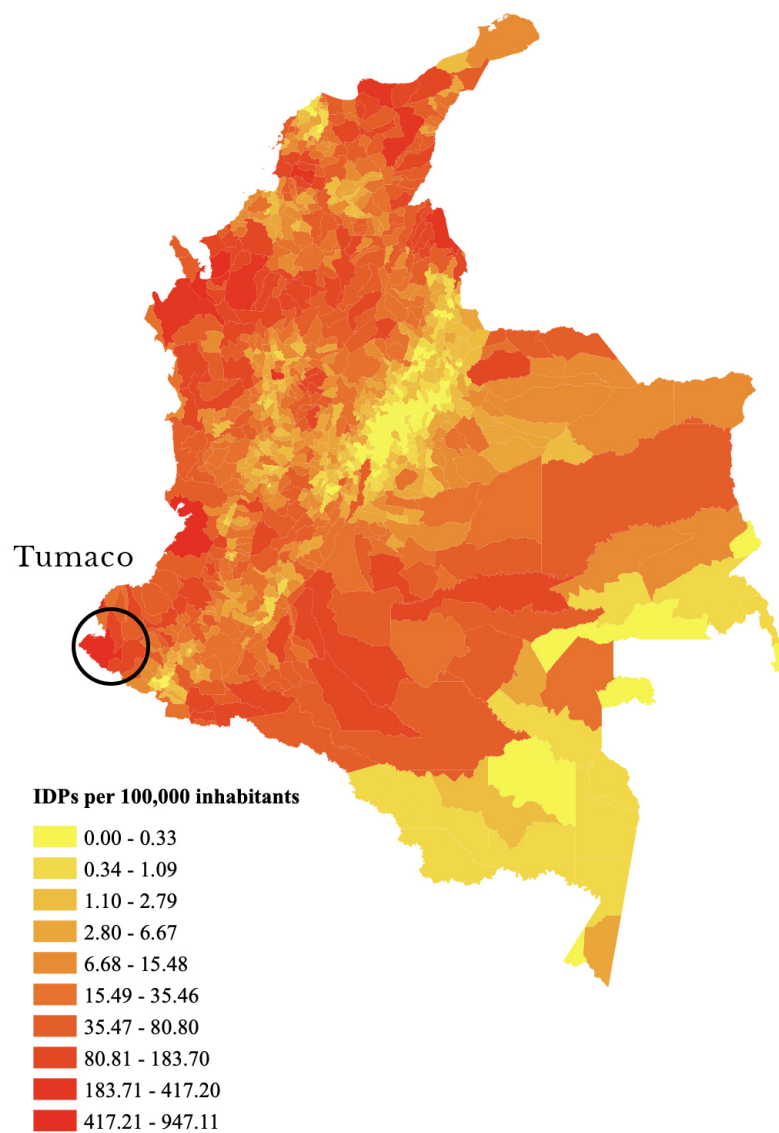
Notes: The map illustrates the intensity of forced displacement by municipality —number of IDPs per 100,000 inhabitants— based on data from the Registro Unico de Victimas, retrieved from `https://www.unidadvictimas.gov.co/es/reportes` on December 2021.(Go back to text)

Figure A2: Childcare Centers in Tumaco
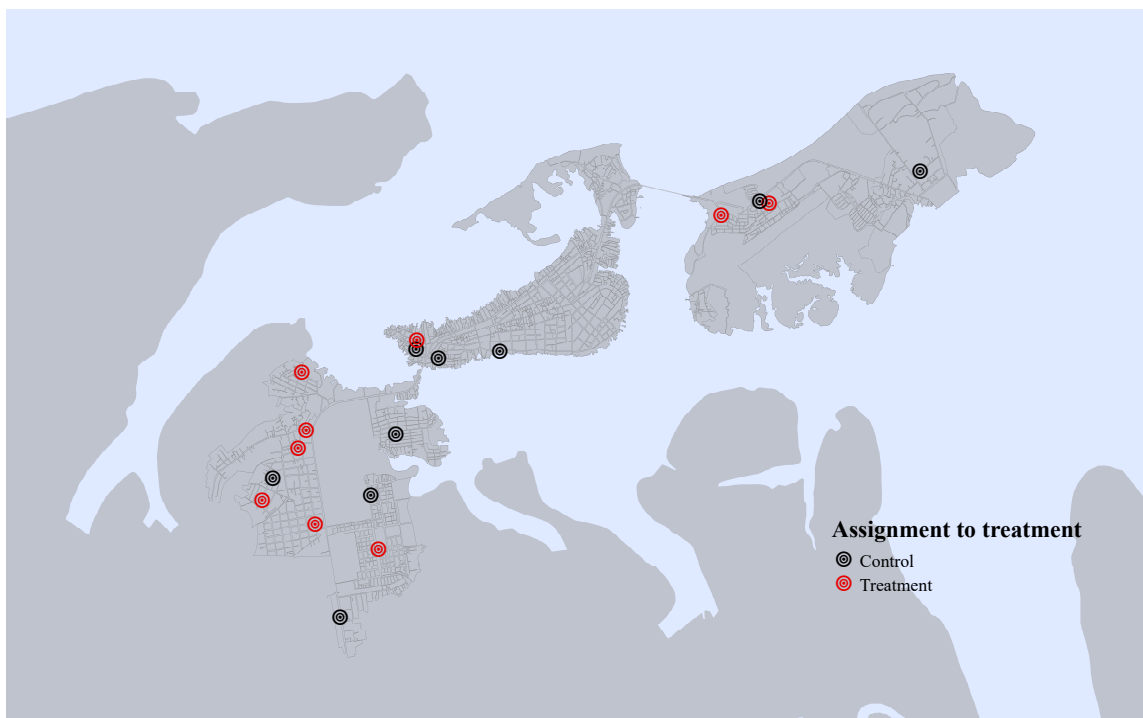
Notes: The map illustrates location of the childcare centers operated by Genesis Foundation, our partner in this study. These center serve approximately 1,600 children, corresponding to 80 percent of children served by public childcare centers in the urban area of Tumaco. Black icons denote centers assigned to the control arm of our experiment, while red icons denote centers assigned to the treatment arm. (Go back to text)

Figure A3: Consort Diagram

19 Early Childhood Development Centers (ECDC) in Urban Area

1 ECDC excluded
♦ Infrastructure was inappropriate + ECDC population did not live in the neighborhood

18 ECDC randomized into intervention groups

Cluster assignment

9 ECDC's assigned to treatment group: Survey + Semillas de Apego

1139 eligible families

9 ECDC's assigned to control group: Survey

1118 eligible families

Define rollout of groups (number of groups per ECDC) over time

Individual assignment

Randomization of dyads for each cohort

712 dyads enrolled in baseline

Randomization of dyads for each cohort

662 dyads enrolled in baseline

Measurement

Baseline: Before starting intervention
Follow-up 1: At the end of the intervention - Short-term effects
Follow-up 2: 8 months after the end of the intervention - Long-term effects

Notes: The figure illustrates the consort diagram that summarizes the experimental design and the sample selection.(Go back to text)

Figure A4: Participation in the program
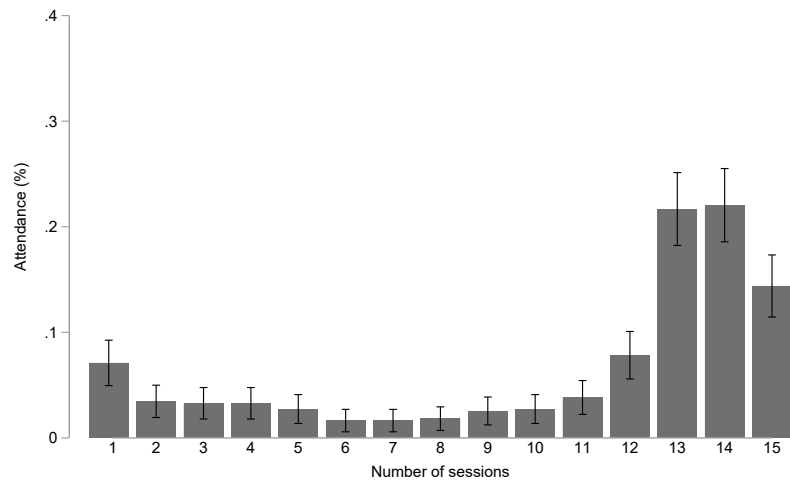
Notes: (Go back to text)

Figure A5: Participation in the program by cohorts
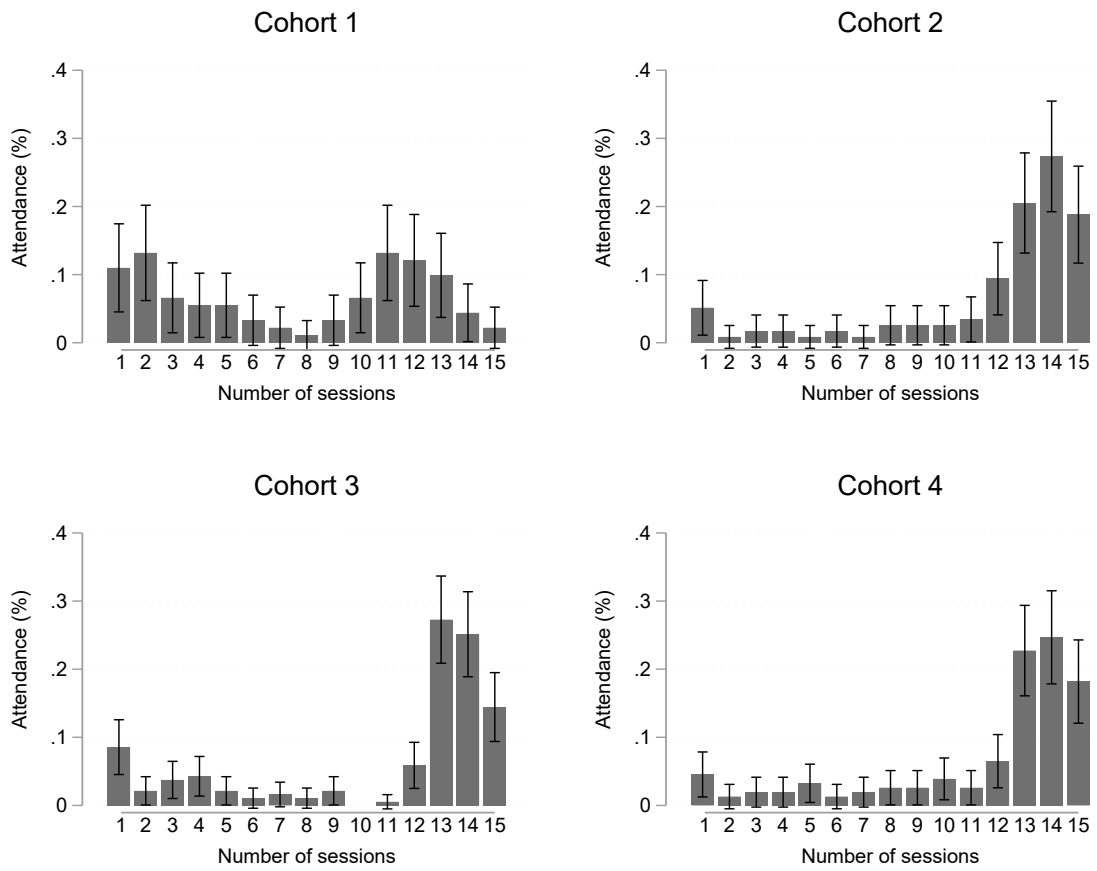
Notes: (Go back to text)

Figure A6: Timeline for data collection

Intervention

BL   *   F1   F2

	BL	*	F1	F2
C1	03/2018	04/2018	08/2018	03/2019
C2	07/2018	08/2018	11/2018	08/2019

Intervention

BL   *   F1   F2

	BL	*	F1	F2
C3	03/2019	04/2019	08/2019	03/2020
C4	07/2019	18/2019	11/2019	08/2020

Notes: (Go back to text)

A6
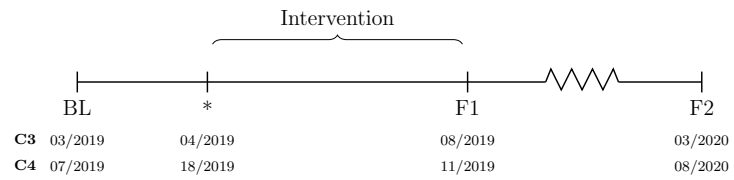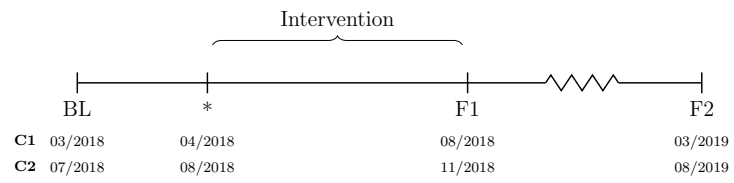
Figure A7: Timeline for data collection and Covid-19 interruption
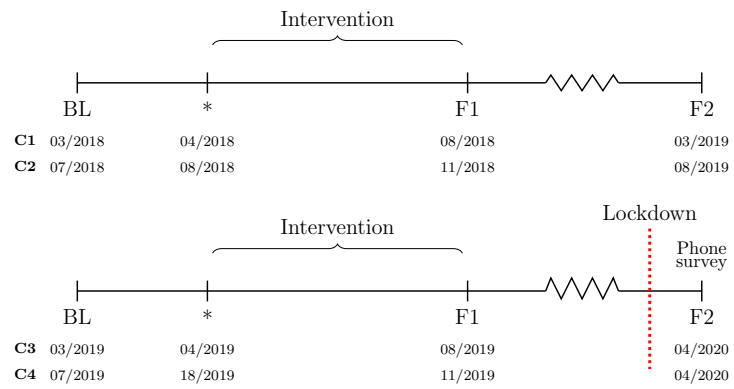
Notes: (Go back to text)

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. (Go back to text)

A8

Table A2: 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
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
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

Notes: (Go back to text)

Table A3: Dosage – Number of Sessions Attended

	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: Conditional attendance refers to the number of sessions attended conditional on attending at least one session. Standard deviations are reported in parentheses. (Go back to text)

Table A4: Selection regression: attendance to 1 session

	1 session
Child age in months	0.00
	[0.002]
Female (=1)	0.04
	[0.029]
Child birth order	-0.00
	[0.016]
Age	0.01*
	[0.002]
Caregiver is the mother (=1)	0.10
	[0.054]
Household size	-0.01
	[0.008]
Number of children under 5yrs	0.06*
	[0.029]
Two-parent hh (=1)	0.11**
	[0.033]
Highest years of education	0.01
	[0.005]
CCT beneficiary (=1)	0.02
	[0.031]
Head is employed (=1)	-0.00
	[0.039]
Asset index	0.01
	[0.014]
Number of violent events	0.01
	[0.009]
IDP (=1)	0.00
	[0.033]
Caregiver Mental Health	-0.04
	[0.023]
Child-Parent Relationships	-0.01
	[0.021]
Child Mental Health	-0.01
	[0.016]
Child Development	-0.01
	[0.021]
Cohort 1	0.00
	[.]
Cohort 2	0.07
	[0.045]
Cohort 3	0.11**
	[0.042]
Cohort 4	0.07
	[0.043]
Constant	-0.15
	[0.152]
Observations	1207
R-squared	0.042
Joint F-statistic	2.687

Notes: Columns 1 shows the estimations of regressing whether the participant attended to at least one session on baseline characteristics. Standard errors reported in brackets, * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A5: 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: Column 1 presents descriptive statistics for the control arm and Column 2 for the treatment arm. Column 3 presents the mean difference between the control and treatment groups. Standard deviations shown in parentheses and standard errors in brakets. * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

Table A6: Selection at Baseline

	Complete Sample (1)	Cohorts1&2 (2)	Cohorts3&4 (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]
Caregiver is the mother (=1)	0.10	0.15*	0.03
	[0.051]	[0.066]	[0.073]
Age	0.00	0.00	0.00
	[0.002]	[0.003]	[0.003]
Household size	-0.01	-0.02	0.00
	[0.008]	[0.012]	[0.011]
Number of children under 5yrs	0.05*	0.09*	0.03
	[0.026]	[0.039]	[0.036]
Two-parent hh (=1)	0.07*	0.13**	0.02
	[0.031]	[0.047]	[0.041]
Highest years of education	0.01	0.00	0.01
	[0.005]	[0.007]	[0.006]
Asset index	0.02	0.05*	0.00
	[0.013]	[0.020]	[0.017]
CCT beneficiary (=1)	0.01	-0.07	0.09*
	[0.029]	[0.045]	[0.038]
Head is employed (=1)	0.01	-0.02	0.04
	[0.036]	[0.053]	[0.050]
Number of violent events	0.02	-0.01	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	-0.04*	-0.06*	-0.03
	[0.016]	[0.024]	[0.020]
Quality of the Relationship	-0.01	-0.01	-0.00
	[0.016]	[0.022]	[0.022]
Style of the Relationship	-0.00	-0.04	0.03
	[0.015]	[0.023]	[0.020]
Child Mental Health	0.03	0.08**	0.00
	[0.017]	[0.025]	[0.021]
Early Childhood Development	-0.02	-0.01	-0.04
	[0.015]	[0.023]	[0.020]
Constant	0.07	0.30	0.00
	[0.142]	[0.225]	[0.181]
Observations	1372	572	800
R-squared	0.030	0.073	0.051
F-test	2.368	3.028	2.475

Notes: Columns 1 to 3 show the estimations of regressing the treatment assignment on baseline characteristics. Column 1 shows the estimation for the complete sample, and columns 2 and 3 restrict the sample to participants assigned to cohorts 1 & 2 and to cohorts 3 & 4 respectively. Standard errors reported in brackets, * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

A13

Table A7: Selection into each Cohort

	Cohort			
	(1)	(2)	(3)	(4)
Child Age in months	-0.005***	0.000	0.004**	0.001
	[0.001]	[0.001]	[0.002]	[0.002]
Female (=1)	0.025	0.036	-0.048*	-0.012
	[0.022]	[0.022]	[0.025]	[0.024]
Child birth order	0.007	0.003	-0.024*	0.013
	[0.011]	[0.012]	[0.013]	[0.013]
Caregiver is the mother (=1)	-0.107**	-0.041	0.138***	0.009
	[0.045]	[0.044]	[0.043]	[0.044]
Age	-0.002	-0.002	0.002	0.002
	[0.002]	[0.002]	[0.002]	[0.002]
Household size	-0.006	-0.006	0.003	0.008
	[0.007]	[0.007]	[0.007]	[0.007]
Number of children under 5yrs	0.045**	0.007	-0.053**	0.002
	[0.022]	[0.023]	[0.023]	[0.025]
Two-parent hh (=1)	-0.009	0.008	-0.015	0.016
	[0.025]	[0.025]	[0.028]	[0.027]
Highest years of education	0.005	-0.006	-0.005	0.006
	[0.004]	[0.004]	[0.004]	[0.004]
Asset index	-0.002	-0.000	0.001	0.002
	[0.011]	[0.011]	[0.012]	[0.011]
CCT beneficiary (=1)	0.087***	0.014	0.022	-0.123***
	[0.023]	[0.024]	[0.027]	[0.025]
Head is employed (=1)	-0.047	0.001	0.031	0.015
	[0.030]	[0.030]	[0.033]	[0.032]
Number of violent events	-0.004	0.005	0.022***	-0.023***
	[0.007]	[0.007]	[0.007]	[0.007]
IDP (=1)	0.019	0.018	-0.113***	0.076***
	[0.024]	[0.025]	[0.028]	[0.027]
Caregiver Mental Health	-0.011	-0.004	0.042***	-0.027*
	[0.013]	[0.013]	[0.014]	[0.015]
Quality of the Relationship	0.035**	0.007	-0.026*	-0.017
	[0.014]	[0.013]	[0.014]	[0.014]
Style of the Relationship	0.042***	-0.012	-0.030**	-0.000
	[0.012]	[0.013]	[0.014]	[0.013]
Child Mental Health	-0.028*	0.039***	-0.018	0.007
	[0.015]	[0.013]	[0.016]	[0.015]
Early Childhood Development	-0.023*	-0.020	0.059***	-0.016
	[0.012]	[0.013]	[0.014]	[0.014]
Observations	1372	1372	1372	1372
Joint Significance F-test	4.714	1.171	4.312	2.669

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, * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

Table A8: Scales

Dimension	Scale	Subscale	Source
Caregiver Mental Health	Questionnaire based on the Symptom Checklist 90-R (SCL-90)	Sensitivity Depression Phobic Anxiety Anxiety Hostility	Derogatis and Unger (2010)
Quality of the Relationship	Parental Stress Index (PSI - short form)	Difficult Child Parent-Child Dysfunctional Interaction Parental Distress	Haskett et al. (2006)
	Child-Parent Relationships (CPR)	Conflicts Closeness Depedence	Driscoll and Pianta (2011)
Style of the Relationship	Interactions	Positive discipline Negative discipline Routines	Bernal et al. (2015)
Child Mental Health	Trauma Symptom Checklist for Young Children (TSCYC)	Anger Anxiety Sexual concern Depression Posttraumatic Stress Dissociation	Briere et al. (2001)
Early Childhood Development	International Development Early Learning Assessment (IDELA)	Numeracy Literacy Socio-Emotional Motor Executive functions	Pisani et al. (2018)
	Brief Toddler Socioemotional Assessment (BITSEA)	Socio-Emotional / Behavioral Problems Delays in Social-Emotional Competences	Briggs-Gowan and Carter (2007)

Notes: (Go back to text)

Table A9: Psychometrics – Composite Indices

	Cronbach's alpha	RMSEA
Caregiver's Mental Health	0.952	0.059
Quality of the Relationship	0.901	0.059
Style of the Relationship	0.476	0.077
Child Mental Health	0.921	0.054
Child Development	0.739	0.056

Notes: All composite indices were measured at baseline for the entire sample, Child Development and Quality of the Relationship were measured at the 8-month and 1-month follow-ups, respectively, only for the control arm.(Go back to text)

Table A10: Psychometrics: Individual Scales

	Cronbach's alpha	RMSEA
Caregiver's Mental Health		
SCL	0.952	0.059
Quality of the Relationship		
PSI	0.882	0.068
CPR	0.694	0.083
Style of the Relationship		
Interactions	0.476	0.077
Child Mental Health		
TSCYC	0.921	0.054
Child Development		
BISTEA	0.756	0.050
IDELA	0.871	0.062

Notes: All individual scales were measured at baseline for the entire sample with the exceptions of the CPR scale, which was measured at the 1-month follow-up for the control arm, and the IDELA scale, which was measured at the 8-month follow-up for the control arm.(Go back to text)

Table A11: 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

Notes: (Go back to text)

Table A12: Balance: Treatment Vs Control for Non-Attritor Sample

	(1) Total	(2) Control	(3) Treatment	(4) Diff (2)-(3)
Panel A. Child				
Child Age in months	35.21	34.90	35.50	-0.60
	(8.346)	(8.205)	(8.470)	[0.459]
Female (=1)	0.50	0.48	0.51	-0.03
	(0.500)	(0.500)	(0.500)	[0.028]
Child birth order	1.90	1.88	1.93	-0.05
	(1.186)	(1.175)	(1.196)	[0.065]
Panel B. Caregiver				
Caregiver is the mother (=1)	0.87	0.85	0.88	-0.03
	(0.338)	(0.356)	(0.320)	[0.019]
Female (=1)	0.96	0.94	0.97	-0.04**
	(0.203)	(0.241)	(0.160)	[0.011]
Age	29.11	28.93	29.29	-0.36
	(9.335)	(9.146)	(9.511)	[0.514]
Years of education	11.89	11.90	11.88	0.02
	(3.658)	(3.555)	(3.753)	[0.201]
Panel C. Household				
Household size	5.01	4.99	5.03	-0.04
	(2.009)	(1.958)	(2.056)	[0.111]
Number of children under 5yrs	1.29	1.26	1.32	-0.06
	(0.558)	(0.494)	(0.610)	[0.030]
Two-parent hh (=1)	0.71	0.68	0.74	-0.06*
	(0.454)	(0.467)	(0.440)	[0.025]
Highest years of education	12.74	12.60	12.87	-0.26
	(3.297)	(3.323)	(3.269)	[0.182]
Asset index	-1.65	-1.71	-1.59	-0.11
	(1.193)	(1.184)	(1.200)	[0.066]
Hh income per capita	241.01	246.90	235.56	11.35
	(313.976)	(338.466)	(289.655)	[17.420]
CCT beneficiary (=1)	0.43	0.42	0.44	-0.02
	(0.495)	(0.494)	(0.496)	[0.027]
Head is employed (=1)	0.82	0.82	0.83	-0.00
	(0.380)	(0.382)	(0.379)	[0.021]
Head has formal job (=1)	0.21	0.19	0.22	-0.03
	(0.407)	(0.394)	(0.418)	[0.022]
Days hh head works in a month	23.72	23.51	23.92	-0.41
	(6.470)	(6.597)	(6.353)	[0.400]
Panel D. Violence				
Victim of direct violence (=1)	0.82	0.80	0.84	-0.04
	(0.384)	(0.400)	(0.367)	[0.021]
Number of violent events	2.38	2.25	2.49	-0.24*
	(1.798)	(1.773)	(1.813)	[0.099]
IDP (=1)	0.57	0.57	0.58	-0.01
	(0.495)	(0.496)	(0.494)	[0.027]
Panel E. Indices				
Caregiver Mental Health	-0.08	0.01	-0.15	0.16**
	(1.058)	(0.995)	(1.108)	[0.058]
Quality of the Relationship	-0.03	-0.01	-0.05	0.04
	(1.030)	(1.004)	(1.054)	[0.057]
Style of the Relationship	-0.01	-0.00	-0.02	0.02
	(0.964)	(0.998)	(0.931)	[0.053]
Child Mental Health	-0.01	-0.01	-0.01	0.00
	(0.997)	(1.004)	(0.992)	[0.055]
Early Childhood Development	-0.03	0.00	-0.07	0.08
	(1.056)	(0.994)	(1.109)	[0.058]
Observations	1318	633	685	1318
Joint F-statistic	2.398			

Notes: Column 1 presents descriptive statistics for the entire sample. Column 2 and 3 corespond to the control and treatment arms, respectively. Column 4 presents the mean difference between the control and treatment groups. Standard deviations shown in parentheses and standard errors in brakets. * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A13: Attrition

	1-month		8-month	
	(1)	(2)	(3)	(4)
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

Notes: Columns 1 to 4 show the estimations of regressing whether the participant is an attritor at the 1-month or 8-month followups respectively, on baseline characteristics. Columns 2 and 4 additionally control for cohort fixed effects. Standard errors reported in brackets, * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A14: Impact on Caregivers' Mental Health by Subdimensions

	ITT		% at risk	
	1-month	8-month	1-month	8-month
Anxiety	-0.007	0.081	0.008	-0.002
	[0.040]	[0.082]	[0.015]	[0.023]
Depression	0.015	0.015	-0.004	-0.004
	[0.051]	[0.089]	[0.018]	[0.035]
Phobic Anxiety	0.003	0.152*	-0.006	-0.059*
	[0.043]	[0.087]	[0.017]	[0.031]
Sensitivity	-0.032	0.159*	0.026	-0.059**
	[0.054]	[0.087]	[0.017]	[0.024]
Hostility	0.020	0.181**	0.009	-0.035
	[0.047]	[0.071]	[0.013]	[0.021]
Observations	1316	522	1317	522
Controls	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes

Notes: Results of the lagged dependent variable model on the Caregivers' Mental Health subdimensions, measured at the 1-month and 8-month followup respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order,child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. % at risk is an indicator variable that takes the value of 1 if the symptoms are at least one standard deviation below the baseline sample mean. Cluster robust standard errors are reported in brackets (CR2).; * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A15: Results at 1-month followup - 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 and 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).; * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A16: 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 and 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).; * ρ <0.1. ** ρ <0.05. *** ρ <0.01. (Go back to text)

Table A17: 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	
Quality of the Relationship	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	
Style of the Relationship	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	

Notes: Results of the lagged dependent variable model where the outcome indicated in each panel is measured at the 1-month and 8-month followup respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. Columns 3 and 6 are restricted to the Internally Displaced Population (IDP) subsample at baseline while columns 2 and 5 are restricted to the non-IDP subsample at baseline. Columns 4 and 7 report the p-value of the difference between the two coefficients. Standard errors are reported in brackets. * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A18: 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	
Quality of the Relationship	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	
Style of the Relationship	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	

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. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include the caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. The child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. Columns 3 and 6 are restricted to the subsample of caregivers whose baseline mental health was above the at-risk threshold while columns 2 and 5 are restricted to the not at-risk subsample. Columns 4 and 7 report the p-value of the difference between the two coefficients. Standard errors are reported in brackets. * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A19: 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	
Quality of the Relationship	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	
Style of the Relationship	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	

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. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order,child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. Columns 3 and 6 are restricted to the subsample whose asset index at baseline was below the median while columns 2 and 5 are restricted to the subsample whose asset index at baseline was above the median. Columns 4 and 7 report the p-value of the difference between the two coefficients. Standard errors are reported in brackets. * ρ <0.1. ** ρ <0.05. *** ρ <0.01. (Go back to text)

Table A20: Heterogeneity: Child age more or less than three years old

	1-month followup			8-month followup		
	Age (≤ 3)	Age (> 3)	pvalue	Age (≤ 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	
Quality of the Relationship	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	
Style of the Relationship	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	

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. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. Columns 3 and 6 are restricted to the subsample whose child was more than three years old at baseline while columns 2 and 5 are restricted to the subsample whose child was less or equal to three years old at baseline. Columns 4 and 7 report the p-value of the difference between the two coefficients. Standard errors are reported in brackets. * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

Table A21: Impact on the Quality of the Relationship

	ITT		% at risk	
	1-month	8-month	1-month	8-month
Panel. A Parenting Stress Index				
Global Scale	0.014	0.215*	-0.058*	-0.096**
	[0.069]	[0.116]	[0.030]	[0.038]
Difficult Child	0.089	0.284***	-0.071***	-0.107***
	[0.061]	[0.104]	[0.027]	[0.030]
Parent-Child difficult inter.	0.060	0.152	-0.049	-0.062
	[0.071]	[0.095]	[0.031]	[0.039]
Parental distress	-0.106	0.088	0.021	-0.055
	[0.070]	[0.108]	[0.032]	[0.053]
Panel. B Child-Parent Relationship				
Global Scale	0.069	0.048	-0.031	-0.026
	[0.050]	[0.050]	[0.022]	[0.022]
Emotional closeness	0.104***	0.093**	-0.041***	-0.039***
	[0.037]	[0.037]	[0.014]	[0.014]
Conflicts	0.019	0.001	-0.032*	-0.027
	[0.054]	[0.051]	[0.018]	[0.019]
Dependence	0.009	0.004	0.020	0.021
	[0.057]	[0.058]	[0.024]	[0.024]
Observations	1314	1316	1314	1316
Controls	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes

Notes: Results of the lagged dependent variable model on the Relationships' Quality subdimensions, measured at the 1-month and 8-month follow-up respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include the caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. The child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. % at risk is an indicator variable that takes the value of 1 if the symptoms are at least one standard deviation below the baseline sample mean. Cluster robust standard errors are reported in brackets (CR2).; * $\rho < 0.1$. ** $\rho < 0.05$. *** $\rho < 0.01$. (Go back to text)

Table A22: Impact on Relationship Type of Interactions

	ITT	
	1-month	8-month
Positive discipline	0.179***	0.099
	[0.055]	[0.079]
Negative discipline	0.091*	0.091
	[0.053]	[0.078]
Routines	0.122***	0.076
	[0.047]	[0.086]
Observations	1316	522
Controls	Yes	Yes
Cohort Fixed Effects	Yes	Yes

Notes: Notes: Results of the lagged dependent variable model on the Relationships Type of Interactions subdimensions, measured at the 1-month and 8-month followup respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order,child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. Cluster robust standard errors are reported in brackets (CR2).; * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

A29

Table A23: Impact on Child Mental Health Subdimensions

	ITT		% at risk	
	1-month	8-month	1-month	8-month
Anger	0.055	0.121*	-0.007	-0.066**
	[0.049]	[0.070]	[0.020]	[0.028]
Anxiety	0.106**	0.077	-0.037**	-0.092**
	[0.042]	[0.094]	[0.017]	[0.036]
Depression	0.050	0.220**	-0.024	-0.082**
	[0.062]	[0.089]	[0.020]	[0.036]
Dissociation	0.004	0.060	-0.009	-0.073***
	[0.065]	[0.064]	[0.020]	[0.021]
Sex Concern	-0.005	0.122*	-0.009	-0.067**
	[0.048]	[0.065]	[0.017]	[0.033]
Post Traumatic Stress	0.018	0.126	-0.025	-0.089**
	[0.065]	[0.094]	[0.030]	[0.037]
Observations	1316	522	1317	522
Controls	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes

Notes: Results of the lagged dependent variable model on the Child Mental Health subdimensions, measured at the 1-month and 8-month followup respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. % at risk is an indicator variable that takes the value of 1 if the symptoms are at least one standard deviation below the baseline sample mean. Cluster robust standard errors are reported in brackets (CR2).; * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

Table A24: Impact on Child Development Subdimensions

	ITT		% at risk	
	1-month	8-month	1-month	8-month
Panel. A BITSEA	0.144***	-0.028	0.227**	-0.089***
	[0.043]	[0.020]	[0.101]	[0.033]
Socioemotional	0.120**	0.039	-0.012	-0.064*
	[0.047]	[0.108]	[0.017]	[0.038]
Behavioural problems	0.068	0.250***	-0.016	-0.066***
	[0.054]	[0.081]	[0.019]	[0.026]
Panel. B IDELA		-0.078		-0.005
		[0.125]		[0.042]
Motor		0.176		-0.028
		[0.125]		[0.031]
Literacy		-0.002		0.006
		[0.074]		[0.033]
Numeracy		0.114		0.018
		[0.116]		[0.035]
Socioemotional		0.218		0.024
		[0.263]		[0.035]
Executive functions		0.295*		0.005
		[0.171]		[0.024]
Observations	1376	506	1376	506
Controls	Yes	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes	Yes

Notes: Results of the lagged dependent variable model on the Child Development subdimensions, measured at the 1-month and 8-month followup respectively. The outcome is regressed on treatment assignment, the lagged-dependent variable at baseline, a cohort fixed-effect, and a set of controls in the baseline. Controls include caregiver's age, asset index, caregiver's years of education, 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, child's sex, the child's birth order, child's weight for height Z-score, child age in days, and the caregiver's mental health. Child's age in days is excluded as a control for Child Development since the index in already standardized with respect to the Child's age. % at risk is an indicator variable that takes the value of 1 if the symptoms are at least one standard deviation below the baseline sample mean. Cluster robust standard errors are reported in brackets (CR2).; * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$. (Go back to text)

Table A25: Selection into each Cohort

Variable	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Child Age in months	0.092	-0.195	0.017	0.261
	(0.447)	(0.093)	(0.859)	(0.012)
Female (=1)	0.140	0.024	-0.012	0.152
	(0.244)	(0.837)	(0.908)	(0.142)
Child birth order	-0.026	-0.218	0.092	0.193
	(0.833)	(0.060)	(0.346)	(0.062)
Mother is caregiver (=1)	0.193	0.057	-0.007	0.090
	(0.110)	(0.632)	(0.937)	(0.393)
Female (=1)	0.154	0.215	0.120	0.166
	(0.199)	(0.067)	(0.203)	(0.102)
Age	-0.053	0.026	0.174	-0.040
	(0.661)	(0.821)	(0.074)	(0.703)
Years of education	0.183	-0.089	-0.010	-0.076
	(0.128)	(0.443)	(0.914)	(0.468)
Household size	-0.043	-0.076	0.028	0.152
	(0.720)	(0.515)	(0.770)	(0.141)
Number of children under 5yrs	0.170	0.110	-0.152	0.300
	(0.158)	(0.345)	(0.116)	(0.004)
Two-parent hh (=1)	0.056	0.336	0.118	-0.042
	(0.638)	(0.004)	(0.226)	(0.684)
Highest years of education	0.259	0.026	0.107	-0.003
	(0.031)	(0.823)	(0.272)	(0.973)
Asset index	0.231	0.195	0.084	-0.103
	(0.055)	(0.094)	(0.390)	(0.318)
Hh income per capita	0.033	-0.041	0.056	-0.184
	(0.786)	(0.724)	(0.564)	(0.075)
CCT beneficiary (=1)	-0.134	-0.173	0.198	0.159
	(0.267)	(0.139)	(0.040)	(0.127)
Head is employed (=1)	0.172	-0.251	0.000	0.058
	(0.153)	(0.032)	(0.998)	(0.562)
Head has formal job (=1)	0.060	0.085	0.243	-0.075
	(0.621)	(0.462)	(0.012)	(0.482)
Victim of direct violence (=1)	0.176	-0.145	0.273	0.111
	(0.144)	(0.213)	(0.005)	(0.277)
Number of violent events	0.045	-0.111	0.377	0.163
	(0.708)	(0.342)	(0.000)	(0.116)
IDP (=1)	0.041	-0.118	0.192	-0.002
	(0.735)	(0.307)	(0.047)	(0.987)
Observations	277	296	429	374

Notes: Columns 1 to 4 show the standardized differences between the treatment and control group means in each cohort. p-values are reported in parentheses, * $\rho <0.1$. ** $\rho <0.05$. *** $\rho <0.01$.