DISCUSSION PAPER SERIES

Experimental Impacts of a Virtual Parenting Program with Mothers and Fathers

Lara Lebedinski ®
Pedro Carneiro ®
Tamara Arnold Urzua ®
Julie Perng ®
Ana Maria Munoz Boudet
Katia Herrera Sosa

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

Experimental Impacts of a Virtual Parenting Program with Mothers and Fathers

Lara Lebedinski ®
University of Vienna

Pedro Carneiro ®
University College London, CEMMAP, IFS and IZA

Tamara Arnold Urzua ®
The World Bank Group

Julie Perng ®
University of Göttingen

Ana Maria Munoz Boudet
The World Bank Group

Katia Herrera Sosa
The World Bank Group

ABSTRACT

Experimental Impacts of a Virtual Parenting Program with Mothers and Fathers*

In this paper we evaluate experimentally the impacts of a parenting program delivered virtually to 1,431 families with children ages 2-6 years old in Serbia. We compare two program modalities to a control group. In the first (standard) modality, only the main caregiver -mainly mothers- participates in the training, while in the second (plus), two caregivers -mothers and fathers- participate in the training. We find that the standard intervention has a positive effect of 0.28 standard deviations on parent-reported child development outcomes, but no such impact is found for the plus modality. We fail to observe statistically significant impacts of either treatment across most variables measuring parental behaviors and home environments, with two exceptions: parents in the standard treatment became more conscious about their child's learning, while parents in the plus modality became less likely to use physical punishment to discipline children.

Corresponding author:
Pedro Carneiro
Department of Economics
University College London
Gower Street
London WC1E 6BT
United Kingdom

E-mail: p.carneiro@ucl.ac.uk

1. Introduction

Early childhood environments are central to childhood development (e.g., Behrman et al., 2004; Martinez et al., 2012). Since children spend most of the early years at home with their families, parenting education programs aim to increase school readiness, cognitive development, motor skills and interest in learning, among other outcomes (e.g., Black et al., 2017; Britto et al., 2017; Engle et al., 2007, 2011).

There are different types of parenting programs, from home visits to group parenting sessions, requiring very different types of delivery infrastructure. Some small-scale programs, such as the Jamaica intervention of McGregor et al (1991) have shown how such interventions can produce dramatic changes in the lives of poor children in the long run. However, in spite of their promise, scaling up these programs has been very challenging (Attanasio et al, 2022).

One fairly untested way of providing these programs on a large scale is to do it virtually. Virtual provision began to be taken more seriously during the COVID-19 pandemic, which posed significant problems for the delivery of public services, and the response to these fostered the development of new ways of interacting with the public. It is widely recognized that the switch to virtual interactions brought many opportunities (Kizilcec, et al., 2020; Peters, et al., 2022). They have been effective in many settings, and several virtual programs have remained in place even as the pandemic subsides, due to their potential to reach several users at a low cost (Means, Toyama, Murphy, Bakia, & Jones, 2009; Amaral, Dinarte-Diaz, Dominguez, & Perez-Vincent, 2024; Fordis, et al., 2005).

In this paper we experimentally evaluate a virtual parenting program delivered in two phases, or rounds, in 13 Serbian districts from December 2020 through May 2021. The original program was designed to be implemented face-to-face, but due to the pandemic, the team decided to test the effectiveness of virtually-delivered early childhood development interventions. The program was adapted to be delivered virtually during the COVID-19 pandemic, dramatically reducing its cost, although potentially also affecting its effectiveness.

We study two versions of the program. The *standard* version requires the participation of only one caregiver in the program, typically the mother. The *plus* version of the program required the participation of two caregivers, typically father and mother (and henceforth referred to as such), and included content directed to male caregivers to promote their active engagement with the children. The *plus* version was motivated by the idea that it is important to study whether and how we could increase fathers or male caregivers engagement in child development and early stimulation, especially in a context where women spend twice longer on unpaid work (four hours and 09 minutes per day) than men do, most of which is dedicated to home production and childcare.[3] The program was implemented by public preschool teachers as part of the efforts by the government to improve access and coverage of early childhood education and care.

We analyzed data from 1,431 families with children ages 2-6 years old across 13 districts and 21 municipalities in Serbia. These families were randomly assigned to one of three groups. In the standard treatment (n=365), the main caregiver (most frequently the mother) was invited to participate in a program consisting of eight weekly virtual group sessions, weekly text messages (SMS) and audio messages/phone calls, and received supplemental materials such as parenting guides and activity books as part of the program. In the plus treatment (n=380), both female and male caregivers of a child (most

---

[3]A number that has remain almost unchanged for the last 6 years. Statistical Office of the Republic of Serbia (2023).

frequently mother and father) were invited to participate in the same eight weekly group sessions as in the standard and received the same supplemental materials. In this case both caregivers received weekly text messages, audio messages/phone calls, and the male caregiver received an extra weekly actionable text message. In the Control group (n=686), participants were placed on a waitlist, and were invited to participate in the standard version of the program after the endline data collection was completed. To assess impacts, we collected endline data via a phone interview from August through October 2021. The survey instrument measured parental practices, stress, beliefs and attitudes, as well as child development as reported by the caregiver.

Relative to most studies of parenting programs, this research design is innovative in two main ways: 1) by providing evidence about the effectiveness of a light-touch, virtually-run, publicly delivered parenting program; and 2) by providing evidence about the relative effectiveness of improving parental engagement and childhood outcomes using an intervention focused on the main caregiver (primarily female) versus another intervention requiring additional engagement from the male caregiver.

We find a positive effect of .28 standard deviations of the standard treatment on parent-reported child development outcomes, relative to children in the control group. This finding is robust to multiple hypothesis testing, and to multiple imputation for missing data. We find no evidence of impacts of the plus treatment on these outcomes.

We fail to observe statistically significant impacts of either treatment across most variables measuring parental behaviors and home environments, with two exceptions. Parents in the standard treatment were less likely to agree with the statement that they do not have to worry about their child's learning, as their children will learn everything they need to learn in school. Parents in the plus treatment were significantly less likely than parents in the control group to report physical disciplining their children -shake, spank, hit, slap (with bare hands or an object)- during the month preceding data collection.

We also do not detect an effect of the program on attitudes regarding affection and attachment, the Home Observation Measurement of the Environment score, beliefs that parents (or parents and schools) were the ones most responsible for social or academic skills, number of activities parents do with their children, or parental stress.

It seems surprising that the plus treatment did not show impacts on child development, and also that child development improved despite minimal reported changes in parental behaviors and home environments were minimal. In the results section, we provide some potential hypotheses around the theory of change for these results, and also discuss (and mostly rule out) potential limitations in our survey instrument.

Our paper contributes to the large literature documenting that the early years are a crucial time in a child's life for promoting and laying the foundation for reaching their full developmental potential. The stimulation, care, support and affection that caregivers provide to children is a protective factor and foster children's health, language, cognitive and socioemotional development (Frongillo et al., 2017; Mo et al., 2014; Richardson et al., 2021). Evidence indicates that investing in the first years has ripple effects over the life span of individuals, and can increase their productivity into adulthood (Behrman et al., 2004; Heckman, 2008; Martinez et al., 2012; Naudeau et al., 2011; Sayre et al., 2015).

Research suggests that parents in low-income households may participate in children's learning at lower rates (Repetti et al., 2002; Rubio-Codina et al., 2015), with early stimulation being more important for these children and those with no access to early education. This will be particularly important for children below the minimum mandatory age of preschool education, who are likely to spend most of their time at home.

Since caregivers' knowledge of development plays a key role in the care they can provide, many interventions promoting successful child development have been directed towards caregivers through a variety of program types and intensities (for an overview see Richardson et al., 2021). What these programs have in common is the provision of information and practical activities for parents to undertake, under the assumption that these type of information interventions should help improve parental skills and behaviors. There is evidence showing that this can actually be the case: for instance, messages related to (targeted) parental advice and returns to education can increase the time spent on child development by as much as 24 percent (Cunha et al., 2013; Gandy et al., 2016; Jensen, 2010; York et al., 2019).

The method of information dissemination and intensity of parental engagement programs can vary. There are individualized programs such as home visits like the ones implemented in Jamaica and more recently replicated in South Asia and South America, showing a range of substantive short and long term outcomes or more muted results depending on the setting (Attanasio et al., 2014; Gertler et al., 2014; Grantham-McGregor et al., 1991; Hamadani et al., 2006; Nahar et al., 2012; Vazir et al., 2013). There have also been group-based interventions that have lead to changes in parental behavior and child outcomes (Aboud & Akhter, 2011; Aboud & Yousafzai, 2015; Al-Hassan & Lansford, 2011; Carneiro et al., 2022; Fernald et al., 2017; Kagitcibasi et al., 2009; Singla et al., 2015; Skrypnek & Charchun, 2011; Yousafzai et al., 2014), and combinations of both. For example, Gowani et al. (2014) tests a combination of individual visits with group sessions in Pakistan and finds this model to be effective two years after the intervention ended. In Europe, Huillery et al. (2017) tested a multi-armed trial combining monetary and behavioral incentives to improve kindergarten enrollment in Bulgaria, finding that community meetings improved parental beliefs and aspirations. In Serbia specifically, an evaluation of the "Support, Not Perfection" program found that a group-based program may have increased parenting self-efficacy and decreased depressive symptoms and child behavior problems (Sokolovic et al., 2021).

Lighter-touch interventions (not entirely reliant on visits and meetings) could also improve child development (Chinen & Bos, 2016). For instance, one program in Zambia improved children's reading skills through 3 weekly SMS to parents with short stories for their children to read. Parents were encouraged to set time to listen to the child read the story and answer questions together about the reading. This intervention included monthly parental meetings to discuss and address any issues, which was an important component of the program (Ome & Menendez, 2021). Similarly, Bloomfield et al. (2023), find positive effects on parental investment from an intervention in Uruguay that sent parents three messages a week for 6 months, following an eight-week workshop intervention on early stimulation and nurturing environments. However, Arteaga, Barros, and Ganimian (2024) found that a program with automated phone calls offering parenting advice reduced parental self-efficacy, increased parental anxiety, and had null effects on child development and language.

In some contexts, attending in-person meetings is difficult (due to distance, schedules, or social distancing requirements). If access to and penetration of information and communication technologies (ICT) devices like cell phones is strong, utilizing ICT could be a more feasible channel to disseminate learning resources at-scale. For example, text messaging can be used to provide information, reminders, encouragement,

and reinforcement to promote action to change parental practices (Hashemian et al., 2014; Hurwitz et al., 2015; York et al., 2019; Richardson et al., 2021). It is also an inexpensive and easy-to-use service that allows for brief messages to be sent, read at an individual's convenience, and consulted later on at any time.

These types of light-touch programs have proven successful to improve parental practices, with some also improving child development. A program in Nicaragua tested the impact of sending daily text messages with advice about parenting practices for 10 months. This intervention led to significant changes in parental practices but did not translate in child development outcomes (Barrera et al., 2020). A shorter program in the US proved that receiving text messages for six weeks increased the number of learning activities that caregivers did together with their child. This was particularly true of fathers and parents of boys (Hurwitz et al., 2015). Also in the US, York et al. (2019) document a positive impact on parental involvement at home following eight months of regular text messages targeting the behavioral barriers to engaged parenting, with related positive impacts on child early literacy.

This study contributes to this emerging literature on the use of ICT, such as text messages, to deliver learning materials to parents to encourage child development at home. Additionally, it focuses on a gender component not always tested in other studies. Research has shown that fathers usually spend less time and engage in less activities with children than mothers. However, engaging male caregivers is important for child outcomes, since they make important contributions to child development (Cano et al., 2019; Holmes et al., 2020; Panter-Brick et al., 2014).

There is evidence that father-focused interventions are positive for child development (Rolle et al., 2019), can enhance the relationship between father and child (Magill-Evans et al., 2006) and can impact child outcomes via improvements of men' spousal behavior (Björkman Nyqvist & Jayachandran, 2017; Garcia et al., 2022). However, these programs, largely focusing on increasing the quality of the time men spend with their children, since increasing only the quantity of time does not have a positive effect on children's outcomes, have, for the most part significant – albeit small – effects (Jeong et al., 2016; Holmes et al., 2020).

In the next section we describe the intervention. Section 3 describes the data, section 4 discusses the empirical strategy, section 5 presents the results and in section 6 we conclude.

2. Intervention, Sample, and Evaluation Design

Intervention

Strong from the Start - Give them Wings is a parenting program for caregivers -mostly mothers and fathers- of children aged 2-6. The program, initially developed by the Center for Interactive Pedagogy (CIP), aims to empower parents and caregivers and improve their knowledge, attitudes and behavior to promote early childhood development of children younger than primary school entry age. We teamed with CIP to revise and adapt *Strong from the Start* program design to incorporate lessons from successful parenting programs such as Nobody's Perfect implemented in Canada, or *Nadie es Perfecto* from Chile (Carneiro et al., 2023). Adaptations included shortening the program duration and time demands on parents, both to increase the likelihood of fathers as well as mothers completing the program, including father-specific content, as well as making the program more amenable to be delivered across the country by the existing staff in childcare centers. Following the onset of COVID-19 and its related restrictions, the program was also adapted from a face-to-face delivery modality to a virtual one.

The resulting program consisted of a series of workshops delivered weekly over eight weeks, and covering topics including play, parenting stimulation, family environment, and father involvement, with content pared down to avoid overburdening the families under the new virtual modality.

The virtual delivery mode was developed carefully with the participating preschool institutions to ensure that parents with different levels of access to and knowledge of technology could be reached and participate in the program. It consisted of weekly, facilitated virtual group discussions (a mix of structured knowledge delivery and the sharing of experiences) on Viber.[4] In addition, the facilitator sent to parents a weekly set of pre-prepared SMS and audio messages suggesting activities, calling for reflection, and summarizing discussions.

Workshops were delivered by trained preschool teachers from the preschools located in the selected delivery areas, who were closely monitored and supervised by more experienced educational specialists/preschool teachers (*mentors*). On average, each preschool teacher delivering the program was in charge of a single group of families who went through the program together. The average group size was 6.6 families. [5]

The curricula were semi-structured, and messages were specially designed and standardized to make them understandable, pragmatic, and succinct. The group discussions focused on a new skillset every week, elaborating on potential bottlenecks (and solutions) related to parenting of young children, and provided a post-discussion summary.

In addition, parents received supporting materials especially designed for the program (a program brochure/workbook, children's book, family card games and a package of writing and drawing supplies for children). The materials as well as the intervention were provided in Serbian as well as in Hungarian and Slovak, depending on the language spoken by families.[6]

Two treatment arms were delivered as part of the intervention:

- **Treatment 1 (*standard*): Parenting virtual workshops for main caregiver.** In this treatment group, the main caregiver (generally a female) was invited to participate in the 8 weekly virtual group sessions and received weekly text messages, audio messages/phone calls, and supplemental materials and parental guides.
- **Treatment 2 (*plus*): Parenting virtual workshops for female and male caregivers.** In this treatment group, both mother and father (or main female and male caregiver) were invited to participate in the standard intervention, with the difference that both received the intervention text messages, and the male caregiver received an additional actionable text message a week.

Sample and Evaluation Design

---

[4] Viber is a free voice and instant messaging software application similar to WhatsApp and Telegram. It is widely use in Serbia. As of 2019 Viber was the most used social media application in Serbia (77 percent of respondents used it in the last quarter of 2019 according to Survey on Serbian Digital Habits in 2019, conducted by Kantar, TMG Insights).
[5] Further details, including the average number of families and family groups per preschool institution, can be found in Lebedinski et al., 2024.
[6] The program and supplementary materials received by the families are available upon request.

The study took place in four of the five regions of Serbia (Vojvodina, Belgrade region, Region of South-East Serbia and Region of Sumadia and West Serbia). Participating preschool institutions were located in 13 districts and 21 municipalities.[7]). Within each targeted municipality, the public preschool institution was invited to participate in the program,[OBJ][8][9] and those preschool institutions recruited families with children ages 2-6 within their catchment areas through brochures, pamphlets, and social media. Interested families applied through an online platform.[10]

All families enrolled to participate in the parenting program were randomly assigned into either one of the two treatment arms or the control arm. The randomization was stratified based on income source, education level of parents, and preferred language (Serbian or other). There were 365 families allocated to each of the two treatment arms and 686 families in the control condition. Control families were placed on a waitlist and invited to participate in the virtual parenting workshops for main caregiver, the standard intervention, after the endline data collection was completed.

The lists of families in each treatment arm, along with the parenting group assignment (i.e., the distribution of families into groups for the workshop delivery), were given to the preschool institutions implementing the intervention, each of which had both treatment arms and control families (the individual-level randomization occurs within each preschool institution; the sample was stratified by preschool institution, education level, income, and language). The evaluation team provided suggested classroom groups. The preschools were given freedom to make minor adaptations to the classroom groups within arms, but not to treatment assignments, to accommodate preferences for classroom sizes and the location of parents. They then assigned facilitators to the groups, as well as confirmed attendance, after issuing invitations.

The intervention was implemented in two rounds. The first one included 14 preschools and delivered the program from December 2020 until February 2021. The second round included eight preschools and delivered the program from March to May 2021. In total, 82 groups of six to ten families were formed.[11]

---

[7] See Appendix B for a list of locations.

[8] With very few exceptions, each municipality in Serbia has only one preschool (administrative) institutional unit with several branches. We selected preschool institutions taking into account the following characteristics: municipality's development and regional representation, gap in enrollment of children under 5, the preschool's interest, and involvement with the new preschool curriculum.

[9] In Serbia each local self-government (which are similar to municipalities) has one public preschool institution. There are two exceptions: Niš has two public preschool institutions, both of which are included in our sample. In Belgrade, each of the 17 municipalities has its own preschool institution.

[10] In roughly half of the households who applied, the highest level of education was secondary school (46.0%), and the remaining half of households are distributed between college or higher education (28.9%) and primary school (25.0%) (compared to Serbians aged 30-34, 34% who have completed tertiary education (The World Bank, 2023)). In most households, salaried income is the main source of income; however, 19.27% in the total sample have reported that their main income is social assistance. The share of households with a mobile phone is 97.6%, and 93.0% has internet; this is higher than the general population, for which 94.4% had a mobile phone, and 85.6% had broadband internet connection in 2023 (Statistical Office of the Republic of Serbia, 2023). Most households chose to attend the program in the Serbian language (96.2%), however the program was also delivered in Hungarian (2.9% of households) and Slovak (0.9% of households). (In the 2011 Serbian Census, a little over 83% (3.5%) of Serbians are the Serbian (Hungarian) ethnicity. Based on language, Serbians could potentially be over-represented in our list of participants.)

[11] In Lebedinski et al., 2024, we compare the characteristics of participants and dropouts, and describe the number of families and family groups per preschool institution.

When there were several parents speaking a minority language in a preschool institution, efforts were made to group them together and for them to receive the training and materials in their preferred language. In total, there were two Hungarian speaking groups.[12]

Figure 1 Sample Size

Treatment arms	Phase/round 1	Phase/round 2	Total
Control	530	156	686
Treatment 1 (*standard*)	270	95	365
Treatment 2 (*plus*)	270	110	380
Total	1,070	361	1431

3. Data and Descriptive Statistics

Data Collection

The enrollment form was used as the baseline survey. Data was collected in Serbian. The form was filled out by the interested families (or preschool institution proxies) in November 2020 (first phase) and February 2021 (second phase). It included information about location, education level of parents, main source of income, connectivity, number and age of children, preschool enrollment, and language (see the Appendix document, Section 3 for the full instrument).[13]

The data on outcomes (for both phases), was collected through an endline survey administered to the main caregivers in the treatment and control groups, who had consented to having their contact data shared with the data collection firm.[14] Female caregivers surveyed were also asked to provide the contact information for male caregivers to respond to a shortened version of the survey (Figure 2 below shows the complete timeline).

The survey was administered via phone interview from August through October 2021. It measured parental practices, parental stress, parental beliefs and attitudes, and child development as reported by the caregiver.

Figure 2 Timeline[15]

Oct. 10-Nov. 11, 2020	Dec. 1 2020- Feb. 17 2021	Feb. 5-24, 2021	March 8-May 14, 2021	Aug.-Oct. 2021	Oct.-Dec. 2021
Baseline data collection first phase/round	Program implementation first phase/round	Baseline data collection second phase/round	Program implementation second phase/round	Endline data collection	Program implementation control rollout

---
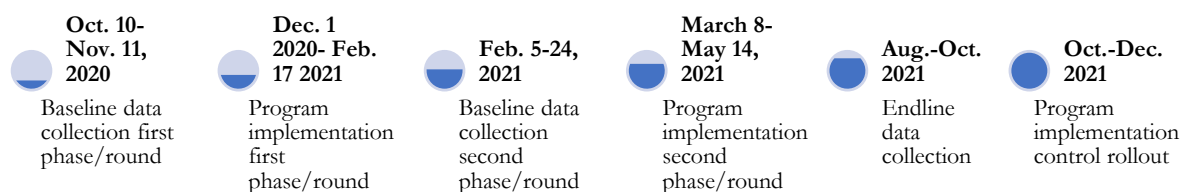
[12] We were unable to form Slovak speaking groups due to the small number of these families; as a result, Slovak-speaking families were assigned to Serbian-speaking groups.

[13] The anonymized data and survey documentation was published within the World Bank Microdata Library under project ID 2259.

[14] Only 87% of families (1,244 families) provided consent to having their data shared with the survey team.

[15] Implementation with the control group started after they were surveyed for the endline data collection

Outcomes

We now turn to the main outcomes we examine. *The nature and quality of the child's home environment* is measured using the Home Observation Measurement of the Environment-Short Form (HOME-SF) instrument (Bradley & Caldwell, 1984; Bureau of Labor Statistics, n.d.). To capture the quality of stimulation at home, this instrument provides an index that covered components such as stimulation, involvement, variety, organization of the home, learning materials, and responsivity. In the construction of this index, all age-appropriate items were included, with the exception of questions that did not apply to the Serbian context. The total score is a simple addition of the dichotomous individual item scores, divided by the number of items of the corresponding age group. The score was standardized within age, using the means and standard deviations in the control group.[16]

Parental engagement is assessed through one of the Family Care Indicators (FCI) (Hamadani et al., 2010; Kariger et al., 2012), which measures the frequency of learning and play activities (such as reading, taking the child out among others) that an adult (mother, father, or another adult) does together with the child. The range of the index score goes between 0-6 activities. In addition, to assess fathers' involvement some questions of HOME-SF were adapted and administered to a subsample of fathers to capture father engagement.

Physical discipline is measured with the discipline module of MICS6 Questionnaire that concerns child disciplining methods towards children aged 0 to 6 years.[17] MICS6 scoring guidelines were followed to create this indicator and assess whether the caregiver shook, spanked, hit, slapped their child in the past month. The total score is a simple addition of the dichotomous individual item scores. The range of the index score goes between 0-1.

*Parental anxiety i*s measured using the Generalized Anxiety Disorder Assessment (GAD-7) which is a seven-item instrument that is used to measure or assess the severity of generalized anxiety disorder. Each item asks the individual to rate the severity of his or her symptoms over the past two weeks. The total score is calculated according to the guidance on GAD, with a higher score indicating higher severity of symptoms.

Parental attitudes and beliefs about parenting and early learning are measured through one module from the Early Parenting Attitudes (EPAQ) questionnaire (Hembacher & Frank, 2020), as well as some questions about attitudes and beliefs about early learning from Silander et al. (2018). The EPAQ module includes 8 items, with a range of scores (averaging scores from each question) from 0 to 6, in which a higher score indicates higher belief in affection and attachment. The questions on attitudes and beliefs are not scored, but rather utilized to summarize beliefs on responsible parties for child development and education.

Child development is measured indirectly through a phone interview with the main caregiver using two instruments, the Early Childhood Development Index (ECDI 2030)[18] and the caregiver report of Anchor Items for Measurement of Early Childhood Development (AIM-ECD) (UNICEF, 2023). Both instruments are similar as they ask the main caregiver of the child to provide information about their child's development with similar wording of questions. These instruments also share the same administration guidelines and

---

[16] The procedure used to recode non-dichotomous responses into a binary variable is available on request.
[17] Questionnaire for Children Under Five https://mics.unicef.org/tools?round=mics6#survey-design
[18] Developed by UNICEF https://data.unicef.org/resources/early-childhood-development-index-2030-ecdi2030/

scoring procedures. Individually, each instrument has demonstrated to be psychometrically robust in low- and middle-income contexts (Pushparatnam A, 2021; United Nations Children's Fund, 2023).

ECDI 2030 is a population-level data collection instrument specifically designed, tested, and validated by UNICEF to be incorporated into large-scale household surveys. It measures three general domains of health, learning and psychosocial well-being of children aged 24 to 59 months (United Nations Children's Fund, 2023). The caregiver report of AIM-ECD was designed, tested, and validated by the World Bank, measuring early literacy, early numeracy, executive function, and socioemotional development of children aged 48 to 83 months (Pushparatnam A, 2021). Using both tools together enables us to cover the whole age range of children in the sample (24 to 59 months), and separate analyses of our data shows that the items and constructs of both instruments have positive correlations and have high reliability.[19] This indicates that the instruments capture very similar constructs and complement each other to promote continuity in the measurement of early childhood development. For our analysis, we use an overall score of child development and standardize it within age, using the means and standard deviations in the control group.

Children and parenting scores vary with the education of the parents. There are two notable exceptions for which there was no significant difference in the answers with respect to parental education. When asked "Who do you think is most responsible for teaching your child academic skills like reading, writing, and mathematics?", irrespective of their educational background, parents agreed that it was either parents alone, or both school and parents. Also, parents of all educational levels believed that in order to bring up, raise, or educate a child properly (in the context of teaching right behaviors or addressing behavior problems), a child does not need to be physically punished.

We also examine whether outcomes vary according to the main source of income of the household. We compare three different sources: (1) main source is income from salary, self-employment, agricultural employment, (2) income from social assistance and (3) income from pension, investment and other sources. Similar to what we found for the education gradient, the answers to the question "Parent or school and parent are both responsible for academic skills" and the questions on physical discipline do not differ by source of income. For all other outcomes, the children from families where the main source of income is a salary, self-employment or agricultural employment have better scores than children in the other two groups. Children from families where the main source of income is "social assistance" and "pension, investment and other" do not differ statistically in their outcomes. In Lebedinski et al. (2024), we provide the parental education (table A1) and income gradients (table A2) for all outcome variables and the correlation matrix of outcome variables (table A3).

Fidelity

Every week, facilitators filled out a short online form describing the topic that was discussed during that week; if they changed anything during the implementation of the session; concerns or questions parents had; challenges encountered during the realization of the session; and additional comments.[20] In addition, they had to identify which families (and whether the male or female caregiver from the family) attended the virtual group meetings, and which families/caregivers actively participated.

---

[19] The Cronbach's alpha for test internal consistency overall is 0.85; for ECDI2030, it is 0.74, while for AIM-ECD, it was 0.78 (Luna-Bazaldua, 2022).

[20] Data was collected via KoboToolBox.

All programed virtual sessions were delivered and all treatment families participating in the program received materials. However, the monitoring data was filled out with minor imperfections. According to the program administrators, all facilitators delivered the scheduled meetings, but some were at times unable to upload their monitoring data due to technical issues with accessing and saving the information using the online form. Overall, 96 percent of the facilitators filled out the monitoring survey for the eight scheduled sessions.

According to the weekly monitoring data, 90 percent of the program was delivered exactly as it was designed (including the scripts, and number and content of the text and audio messages). Most of the facilitators who made some changes reported sharing additional information or adapting the messages slightly. In an audit of facilitator self-reported fidelity to the curriculum and recommended script, we found no significant difference between the two treatment arms (overall, 16 percent of all monitored sessions had mostly small changes to the recommended script that did not affect the integrity or target of the session content).

Program participation was high: nearly all families (95 percent) attended at least one of the weekly virtual meetings delivered, and three out of four families attended all meetings. For the standard treatment, the main caregiver was invited to participate in the meetings, while both caregivers (mother and father) were invited to the plus treatment workshops. Around half of the groups' fathers participated at least one weekly session, with fathers in the standard program participating significantly less (41% versus 65% in the plus). For both treatment groups, mothers attended more meetings than fathers. On average, mothers attended 5.4 out of the 8 meetings while fathers attended only 1.6. Mothers in the standard program attended on average significantly more meetings than those in the plus (5.6 and 5.1, respectively), while fathers in the plus program attended on average significantly more meetings than their counterparts in the standard one (2.3 and 0.9, respectively).

In the paper we focus on the impact of being offered the opportunity to participate in the program on various outcomes. This is an intent to treat parameter, which ignores the differences in dosage or imperfect take-up that were just discussed. We can use the estimates from the previous paragraph to (under some assumptions) obtain the impact of each parenting session (using a standard instrumental variables procedure).

Summary Statistics and Balancing Properties

Table 2 shows that the socio-demographic characteristics of the interviewed parents are balanced across the three treatment groups. As the only exception, in the control group, there is a higher share of unemployed parents looking for work (control: 23 percent, treatment 1: 17 percent and, treatment 2: 16 percent) than in other groups. All other available characteristics do not differ by treatment status.

The primary caregiver in the sample is on average 33 years old, predominantly female (91 percent), and 50 percent of them completed secondary education. At endline, half of the primary caregivers (53 percent) were employed, 19 percent were looking for work, and 27 percent were inactive. For most households in the sample, their main income was derived from wage employment (80 percent). However, 12 percent of the households reported social assistance as their main income source. Almost all households have smartphones (90 percent), but only half of them (49 percent) have a personal computer.

Respondents speak predominantly Serbian at home. Their household's selected random child was on average 4.3 years old at the time of the survey.[21]

Survey attrition

The administrative data at our disposal - information that parents provided when applying for the program – can be used to examine balance across treatment arms. The initial randomization was successful: the socio-economic characteristics by treatment status are balanced at baseline. A higher share of respondents was interviewed at endline from Treatment 1 compared to the control group, while the other comparisons between pairs do not yield statistically significant differences. However, using a joint test we cannot reject that there was no differential attrition by treatment status across the three groups (Table A5a in Lebedinski et al., 2024).

Among the parents who initially applied to the program, 62.2 percent were interviewed at endline. To understand how the baseline sample differs from the endline sample, we examine differences in socio-economic characteristics between interviewed and not interviewed parents at endline (Table A5b in Lebedinski, et al., 2024). A description of dropouts, who were an exceptionally low number (21 out of 745 enrolled), can be found in Lebedinski et al, 2024 (Table A1a).

There are several reasons why parents were not interviewed, and we group them into two large categories. The first group consists of parents who refused to be interviewed. Those are: (1) parents who wanted to (or did) participate in the program but did not give consent to be called for the endline survey (whether explicitly or because they could not be contacted); and (2) parents who gave consent to be contacted for the endline survey but eventually refused to participate when they were contacted for the interview. The second group consists of parents who could not be reached by the survey firm for three main reasons: the parent did not have a phone number, the phone number was not available, or the phone number was incorrect. The first group accounts for 13 percent of the initial sample, or 187 parents. The second group accounts for 24.8 percent of the initial sample.

When examining survey attrition from baseline to endline, we find that parents who were interviewed at endline are of a higher socio-economic status (higher education of respondent, their main income is more likely to be from employment, and they have more assets in terms of a smartphone and/or PC) than parents who were not interviewed. When comparing the two groups of not interviewed parents, we find that those who could not be reached are of lower socio-economic status than parents who refused to participate in the survey.

We further test whether attrition is differentially selective across treatment groups. To do this test, we regress a set of household characteristics on an indicator for attrition, indicators for treatment group, and the interaction of the two, and test if these interactions are equal to zero. We find that there is higher attrition among treated households (both treatment groups) receiving social assistance than in the control group and lower attrition of treated households with a higher socio-economic status compared to control households (lower attrition if households owns a PC for standard treatment and lower attrition if

---

[21] Parents were asked a battery of questions about the development of a specific child from their household. The interviewers were instructed to ask these questions about a child the age of 4 if a child of this age was in the household. If not, interviewers would interview a child closest to the age of four. If there were more children in the household close to 4 years, e.g. one 3 and one 5, the interviewer would randomly select the child to be interviewed. Similarly, in the case of siblings, the interviewer would randomly select the child.

household has a smart phone in the case of treatment plus fathers). These results provide support for differential selection by treatment group.

Since attrition is correlated not only with household characteristics, but also with treatment, we also present estimates which account for selective attrition, based on a multiple imputation procedure (e.g., Rubin, 1988). We find that all but one of our estimated positive impacts – for the variable on agreement about whether or not caregivers worry about their child's learning – are robust to this correction (results with multiple imputation are shown in Lebedinski et al., 2024 (Tables A11-A13)).

The main interview was conducted with 890 primary caregivers. With a goal to better understand how male caregivers responded to the program, respondents were asked for the contact information for secondary caregivers. For all families in the endline, 334 secondary (male) caregivers (38 percent of the sample) were interviewed; not all primary respondents provided secondary contact information. Among the interviewed parents, the baseline characteristics suggest that households where both caregivers were interviewed are of a slightly higher socio-economic status in terms of education of primary caregiver, and in terms of the main income source being wage income. Conditional on having interviewed both caregivers, the baseline characteristics by treatment status are largely balanced (see Section 2 in the Appendix document). There is one marginally significant difference for college education; control parents are of higher education than Treatment parents. As the baseline characteristics are balanced by treatment status, it is possible to look at outcomes for this subsample of parents as a stand-alone sample. Because the response rate for the secondary caregiver survey was relatively low, we give less prominence to these results.

4. Empirical strategy

The impact of the interventions on outcome indicators was estimated using three specifications. We start by comparing the simple difference-in-means between parents' reported child outcomes in the control group (C) to those in the two treatment groups (T1, T2). We then extend this specification by including a set of covariates to improve the precision of our coefficient estimates, as well as to control for any imbalances between groups. Equation (1) shows the full specification, which also controls for preschool institution fixed effects.

$$y_i = \alpha + \beta_1 T1_i + \beta_2 T2_i + \beta_5 X_i + \rho_r + \mu_p + \delta_s + \varepsilon_i \quad (1)$$

y_i is the outcome variable for the family i who is affiliated with preschool institution p. $T1_i$ takes a value of one if the family i was assigned to the parenting virtual workshops for main caregiver (T1), and zero otherwise. Similarly, $T2_i$ takes a value of one if the family i was assigned to the parenting virtual workshops for female and male caregivers (T2), and zero otherwise.

Controls for various individual characteristics are in vector X_i, including gender of primary caregiver (and reported child), education of reporting caregiver, age and preferred language of primary caregiver, household income- and employment-related variables, number of children, and baseline outcomes (as relevant). Strata fixed effects δ_s (controlling for the strata developed during the randomization process) and survey round fixed effects ρ_r (i.e. the phase – first or second - when the preschool institution's treatment groups participated in the program) were also controlled for. Finally, our most complete

specification includes preschool institution fixed effects μ_p. Randomization is at the family level, and μ_p capture all shocks occurring at the level of the preschool, so ε_i are assumed to be i.i.d.

5. Results

Table 3 and Figure 3 show our main results. Overall, parents in Treatment 1 reported a significant and positive impact of parenting training on child development (compared to parents in the control group as well as those in Treatment 2).[22] The combined AIM-ECD and ECDI score was 0.26 to 0.28 standard deviations higher for children in this group than for those in the control. These results are robust to various checks: using the development scores (ECDI + AIM-ECD) for all children regardless of age, restricting the scores to the appropriate age groups of AIM-ECD (ages 4-6) and ECDI (ages 2-4), multiple hypothesis testing, and multiple imputation for missing data.[23] Figure 3 shows the distribution of the child development index by treatment status. The whole distribution of the child development index of the Treatment group 1 is shifted to the right relative to the control and treatment plus groups, with a larger mass in the upper part of the distribution.

Figure 3 Distribution of child development index

---

[22] When running separate regressions on individual components, nine items in six (of seven) domains were significantly positive. Two items in the executive functioning domain showed an improvement of 0.21 to 0.24 standard deviations, one object in the literacy domain showed improvement by 0.23 standard deviations, and three objects in the numeracy domain showed improvement by between 0.16 to 0.22 standard deviations. Treatment 2, on the other hand, showed negative impacts in four objects in four domains: executive functioning and expressive language decreased by over 0.21 standard deviations each. Note that these individual regression findings were not robust to multiple hypothesis testing and conducting domain-specific analysis is not recommended by the creators of the indices due to the small number of items per domain.

[23] Secondary caregivers reported an increase of 0.24 S.D, but the estimate is not statistically significant, we assume that this is the case due to the small sample size (see Table A6 in the Appendix).
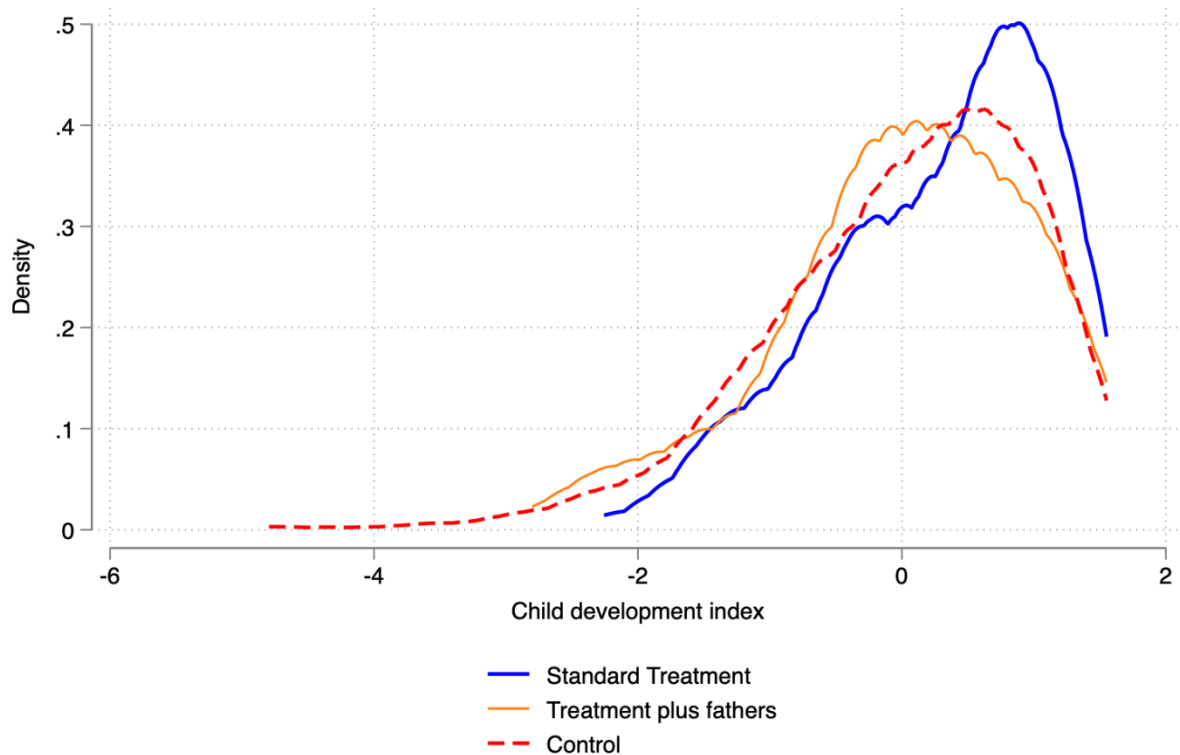
We also looked at stress and beliefs. Table 4 (Column 3) shows that parents in both treatment arms did not have significantly different scores for the GAD-7 index than those in the control group. They also did not hold a significantly different "idea that emotionally close parent-child relationships are important for development" from control parents (Hembacher & Frank, 2020).

Parental beliefs about the process early learning were also collected, and impacts on these outcomes are reported in columns 4-6 of Table 4. Relative to those in the other treatment arms, respondents in the standard treatment group were significantly less likely to agree with the statement that they do not have to worry about their child's learning (because the school is responsible for learning). Note however that this latter finding is not robust to multiple hypothesis testing or to multiple imputation for missing data (Table 5; results on multiple imputation in Lebedinski et al., 2024 (Tables A11-A13)).

Impacts on other parenting beliefs were not different across treatment arms. There was no difference between arms on whether parents were agreed that they were the ones most responsible (versus their child's school or parents and child's school equally) "for teaching [their] child about social skills, like sharing and being patient." Parents were also not significantly different across groups in their views on who was most responsible for academic skills (Column 6 in Table 5).

There is also no significant impact of either treatment arm on beliefs related to physical discipline of children. However, parents in Treatment 2 were up to 0.15 standard deviations less likely to have reported any physical punishment in the past month (Table 4). This indicates that parents who went through Treatment 2 were less likely than parents in the control group to shake, spank, hit, slap (with bare hand

or object) their child in the last month. This finding was robust to multiple imputation for missing data (Table A11 in Lebedinski et al., 2024), but was not robust to multiple hypothesis testing.[24]

Finally, a few measurements were collected to assess the impact of the intervention on how caregivers interacted with their children and provided an environment conducive to learning. The Home Observation Measurement of the Environment-Short Form (HOME-SF) instrument score was not significantly different in either treatment arm.[25] Additionally, male and female caregivers in both Treatment arms were not insignificantly different in the number of activities spent with their children from those in the control group (Table 4).

In sum, our results show that the standard version of the program lead to substantial improvements to child development, in spite of the low cost and apparently low intensity of the intervention. There are, however, two puzzling findings: 1) why are there no impacts of T2 on child development, given that T2 had the same content, albeit with more features concerning father's involvement?[26] 2) how did the impacts of T1 on child development come about, given that T1 produced hardly any observable changes in parental behaviors and home environments?

Regarding question 1, it is possible that the workshops in the plus treatment were different than the ones in the standard intervention. In particular, due to the presence of fathers or male caregivers, these workshops were both larger in size and had a different composition from the standard workshops. This may have led to different workshop experiences in different modalities. For example, if the few fathers present in the workshops disrupted the delivery and understanding of messages about nurturing home environments, or had an impact in the sharing of experiences between participants, or on the formation of support networks among mothers, then it is perhaps plausible that a predominantly-mothers' group works better than one that mixes up mothers and fathers.

It is also possible that the standard intervention motivated the main caregiver to increase attention to both child and caregiver behaviors vis-à-vis the plus version, where such attention might be decreased due to the expectation of a second caregiver also being involved. If these expectations were unmet - fathers in both treatment arms reported an (insignificantly) lower average number of activities with their children than the control group fathers – then this would explain the decreased child development.

Additionally, a review of the evidence shows that parenting programs can increase fathers' knowledge, but not their behaviors (Evans & Jakiela, 2024). Therefore, male caregivers in the plus treatment were potentially unable to provide social learning by instructing and reinforcing teachings that female caregivers did not receive (versus the standard treatment where the female caregivers provided social

---

[24] Additionally, this outcome was not significantly different for secondary caregivers (see Table A9 in Lebedinski et al., 2024).

[25] However, the secondary caregivers' sample, while an imperfect one, provides some support for improvement in the HOME environment score from the workshops: secondary caregivers in Treatment 1 reported a score of .27 standard deviations higher than those in the control group (results on request). In the main caregivers' sample, two items were positive and significant in Treatment 2: whether the caregiver had helped their child learn numbers in the past three days, and whether the family gets together with relatives or friends 2-3 times a month.

[26] We collected data from a subset of male caregivers in addition to the primary survey respondent. The responses of fathers support our main finding of a positive impact of the program on child development in the standard treatment group. Although the coefficient for this group is of similar magnitude as in the case of primary caregiver responses, it is not statistically significant due to low statistical power. In contrast, the coefficient for the plus group is negative and lacks statistical significance; this is aligned with the primary caregiver results.

learning to partners) (Bandura, 1977; Jeong et al., 2019). Unfortunately, we don't have enough information to answer these questions, and they deserve more research.

Regarding question 2, the most likely possibility is that our measures are not rich enough to capture important changes in parental practices that lead to better child outcomes, or they just have too much measurement error. This is perhaps surprising given the wealth of detailed measures - covering multiple aspects of parental attitudes and behaviors (including frequency of some activities) - included in our survey. Note, however, that we do not measure self-efficacy, perceived social support, and have limited data on perception of parental behavior on child development; these are outcomes that Carneiro et al. (2023) found that a large-scale parenting program in Chile had impact on.

Several other studies of large scale and short duration parenting interventions find impacts on parental behaviors and attitudes using similar tools to ours. One would imagine that it is easier to find program impacts on such variables, which are directly targeted by the intervention, than on child development, which is only indirectly targeted. What happens here is the opposite.

A related idea is that the survey instruments that we used were better at capturing the quantity than the quality of parent-child interactions. Because of the pandemic, children and parents in both treatment and control groups were spending almost all their time together, and perhaps for this reason the recorded level of home stimulation and parent-child interactions ended up similar across groups (in fact, the pandemic may have made home environments rather unique to this time period). It is possible that what differed across groups was not so much the quantity but the quality of the interactions, which may have been very different, especially if relationships at home were tense or stressful. Unfortunately, with our data, we cannot really say whether this is what occurred.

Nonetheless, it should also be recognized that, even in many other studies that have used standard mediation analysis, it is often difficult to explain a substantial fraction of the impact of parenting programs on child development through changes in mediators such as parental practices and home environments (Carneiro et al., 2023). Impacts of parenting interventions are often left unexplained. Therefore, it is not unusual to attribute much of the observed change in child outcomes to changes in unobservable parental behaviors. As shown in *Table 5*, findings on T1's impacts on child development – but not other findings related to both treatment arms' impacts on parental behaviors and attitudes were not robust to multiple hypothesis testing.

Heterogeneous effects

We look at heterogeneity with respect to the gender of the child, the age of the child, and the education of the caregiver. These results are shown in tables 10 to 13. In table 10 we report the results for the child development outcomes, and we find some evidence that T2 may have had an impact, but only for boys. We don't find any impacts of child age and parental education on child development. When considering heterogenous impacts on beliefs on affection, learning, social skills, and academic skills, we find only that in the T2 group with college educated parent was the respondent 0.33 standard deviations *less* likely to state that parents don't have to worry about child's learning as school is responsible for it.

For the HOME environment and stress score as measured by GAD, we find that the HOME environment deteriorated by -0.33 standard deviations for older children (aged 5 to 6) in the T2 group and that for male boys in T2 group a reduction in stress of -0.29 standard deviations was reported. Finally, when looking at the outcomes activities with caregivers and physical punishment, we find only an increase in activities

with mother of 0.39 standard deviations in T2 group with college educated parent (Lebedinsk et al., 2024; Appendix B).

Overall, the heterogeneity results show that for certain outcomes and certain subgroups, T2 was beneficial, but there are no obvious patterns of heterogeneity in impacts. We cannot conclude that the program was more successful for some subgroup than for others.

Conclusion

Our research documents the impacts on children and parents of virtual parenting workshops in Serbia. We find that the workshops led to improvements in child development. The AIM-ECD and ECDI score measuring parents' reports on their children's language/cognitive, physical, social-emotional and approaches to learning was significantly higher for children in Treatment 1 (parenting virtual workshops for main caregivers) than for children in the control group.

Additionally, we looked at parenting attitude outcomes. Parents in Treatment 1 were less likely to agree that they did not have to worry about their child's learning (because the school would be responsible) than other parents. While there were no impacts on attitudes related to discipline, parents in Treatment 2 (parenting virtual workshops for female and male caregivers) were significantly less likely than parents in the control group to shake, spank, hit, slap (with bare hand or object) their child in the last month.[27]

Other outcomes were mixed or did not appear to be impacted. For instance, the quantity of parental activities spent with children was overall not significantly improved.[28] Overall home environment (measuring stimulation, variety, and more) was not higher in either treatment arm over the control arm when reported by the primary caregivers, but secondary caregivers in Treatment 1 did indicate higher HOME scores. Additionally, the workshop did not change attitudes on affection and attachment, beliefs in parental responsibility on social or academic skills, or stress.

Our study findings contribute to emerging literature on online and/or low-technology methods of delivering early childhood education information to parents, suggesting that using a combination of SMS messages, booklets, videos, and virtual meetings can be impactful in improving childhood development. In addition, we test a gender component by explicitly encouraging father involvement in one of our study arms – but do not find similar impacts in the variation. Further research into the mechanisms of how Treatment 1 led to improved child outcomes, as well as testing the hypotheses of why Treatment 2 did not lead to similar outcomes, will be beneficial to understanding impactful parenting interventions and gender roles in childhood development.

---

[27] These results were not robust to multiple hypothesis testing.

[28] For primary caregivers with a college education or higher and in Treatment 2, female caregiver activities with children were higher (see Appendix B in Lebedinski et al., 2024).

Bibliography

Aboud, F. E., & Akhter, S. (2011). A cluster-randomized evaluation of a responsive stimulation and feeding intervention in Bangladesh. *Pediatrics*, *127*(5), 1191–1197.

Aboud, F. E., & Yousafzai, A. K. (2015). Global health and development in early childhood. *Annual Review of Psychology*, *66*, 433–457.

Al-Hassan, S. M., & Lansford, J. E. (2011). Evaluation of the better parenting programme in Jordan. *Early Child Development and Care*, *181*(5), 587–598.Attanasio, O., S. Cattan and C. Meghir. (2022). Early Childhood Development, Human Capital, and Poverty. *Annual Reviews of Economics*

Amaral, S., Dinarte-Diaz, L., Dominguez, P., & Perez-Vincent, S.M. (2024). Helping families help themselves: The (Un) intended impacts of a digital parenting program. *Journal of Development Economics*, *166*, 103181.

Arnold Urzua, Carneiro, Herrera Sosa, Lebedinski, Munoz Boudet, & Perng, 2021. "Pre-Analysis plan: Can parenting programs increase their coverage and inclusiveness and still be effective?" https://www.socialscienceregistry.org/versions/102665/docs/version/document

Arteaga, I., de Barros, A., & Ganimian, A. J. (2024). The Challenges of Scaling up Effective Child-Rearing Practices Using Technology in Developing Settings: Experimental Evidence From India.

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

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice Hall: Englewood cliffs.

Barrera, O., Macours, K., Premand, P., & Vakis, R. (2020). *Texting Parents about Early Child Development Behavioral Changes and Unintended Social Effects*. The World Bank Group.

Behrman, J. R., Cheng, Y., & Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *Review of Economics and Statistics*, *86*(1), 108–132.

Björkman Nyqvist, M., & Jayachandran, S. (2017). Mothers care more, but fathers decide: Educating parents about child health in Uganda. *American Economic Review*, *107*(5), 496–500.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., & Lu, C. (2017). & Lancet Early Childhood Development Series Steering Committee. *The Lancet*, *389*(10064), 77–90.

Bloomfield, J., Balsa, A., & Cid, A. (2023). Using behavioral insights in early childhood interventions: the effects of Crianza Positiva e-messaging program on parental investment. *Review of Economics of the Household (21)*, 95-130.

Bureau of Labor Statistics. (n.d.) The HOME (Home Observation Measurement of the Environment). https://www.nlsinfo.org/content/cohorts/nlsy79-children/topical-guide/assessments/home-home-observation-measurement

Bradley, R. H., & Caldwell, B. M. (1984). The HOME Inventory and family demographics. *Developmental Psychology*, *20*(2), 315.

Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., & Vaivada, T. (2017). & Lancet Early Childhood Development Series Steering Committee. *The Lancet*, *389*(10064), 91–102.

Cano, T., Perales, F., & Baxter, J. (2019). A matter of time: Father involvement and child cognitive outcomes. *Journal of Marriage and Family*, *81*(1), 164–184.

Carneiro, P. M., Galasso, E., Lopez Garcia, I. X., Bedregal, P., & Cordero, M. (2019). Impacts of a large-Scale parenting program: Experimental evidence from Chile. *Journal of Political Economy.*

Chinen, M., & Bos, J. (2016). *Impact evaluation of the Save the Children early childhood stimulation program in Bangladesh: Final report*. American Institutes for Research. http://pubdocs.worldbank.org/en/163331484753270396/SIEF-Bangladesh-Endline-Report-Nov2016FINAL.pdf

Carneiro, P., P. Bedregal, M. Cordero, E. Galasso and I. Garcia. (2022). Evaluating the Effect of a Large Scale Parenting Intervention in Chile: Evidence from *Nadie Es Perfecto*. Working paper.

Cunha, F., Elo, I., & Culhane, J. (2013). *Eliciting maternal expectations about the technology of cognitive skill formation* (Issue w19144)). National Bureau of Economic Research.

Engle, P. L., Black, M. M., Behrman, J. R., Mello, M. C., Gertler, P. J., Kapiriri, L., & Group, I. C. D. S. (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*, *369*(9557), 229–242.

Engle, P. L., Fernald, L. C., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., De Mello, M. C., Hidrobo, M., Ulkuer, N., & Ertem, I. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, *378*(9799), 1339–1353.

Evans, D. K., & Jakiela, P. (2024). The Role of Fathers in Promoting Early Childhood Development in Low- and Middle-Income Countries.

Fernald, L. C., Prado, E., Kariger, P., & Raikes, A. (2017). *A toolkit for measuring early childhood development in low and middle-income countries*.

Fordis, M., King, J. E., Ballantyne, C. M., Jones, P. H., Schneider, K. H., Spann, S. J., ... & Greisinger, A. J. (2005). Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomized controlled trial. *jama*, *294*(9), 1043-1051.

Frongillo, E. A., Kulkarni, S., Basnet, S., & Castro, F. (2017). Family care behaviors and early childhood development in low- and middle-income countries. *Journal of Child and Family Studies*, *26*(11), 3036–3044.

Gandy, K., King, K., Streeter Hurle, P., Bustin, C., & Glazebrook, K. (2016). *Poverty and decision-making. How behavioral science can improve opportunity in the UK*. Behavioural Insights Team.

Garcia, I. L., Fernald, L. C., Aboud, F. E., Otieno, R., Alu, E., & Luoto, J. E. (2022). Father involvement and early child development in a low-resource setting. *Social Science & Medicine* (302) 114933.

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

Gowani, S., Yousafzai, A. K., Armstrong, R., & Bhutta, Z. A. (2014). Cost effectiveness of responsive stimulation and nutrition interventions on early child development outcomes in Pakistan. *Annals of the New York Academy of Sciences*, *1308*(1), 149–161.

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

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

Hamadani, J. D., Tofail, F., Hilaly, A., Huda, S. N., Engle, P., & Grantham-McGregor, S. M. (2010). Use of family care indicators and their relationship with child development in Bangladesh. *Journal of Health, Population, and Nutrition*, *28*(1), 23.

Hashemian, T. S., Kritz-Silverstein, D., & Baker, R. (2014). The feasibility and acceptability of a text messaging intervention to improve oral health behavior and knowledge. *Journal of Public Health Dentistry*, *75*, 34–41.

Heckman, J. (2008). The Case for Investing in Disadvantaged Young Children, in Big Ideas for Children: Investing in Our Nation's Future. *Economic Inquiry*, *46*(3), 289–324.

Hembacher, E., & Frank, M. C. (2020). *The Early Parenting Attitudes Questionnaire: Measuring Intuitive Theories of Parenting and Child Development* (Vol. 6, Issue 1). Psychology. https://doi.org/doi:

Holmes, E. K., Egginton, B. R., Hawkins, A. J., Robbins, N. L., & Shafer, K. (2020). Do Responsible Fatherhood Programs Work? A Comprehensive Meta-Analytic Study. *Family Relations*, *69*, 967–982.

Huillery, É., Laat, J., & Gertler, P. (2017). *Supporting Disadvantaged Children to Enter Kindergarten: Experimental evidence from Bulgaria.* Washington, DC: The World Bank.

Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., & Wartella, E. (2015). Supporting Head Start parents: Impact of a text message intervention on parent–child activity engagement. *Early Child Development and Care*, *185*(9), 1373–1389.

Kariger, P., Frongillo, E., Engle, P., Rebello Britto, P., Sywulka, S. and Menon, P. 2012. Indicators of Family Care for Development for Use in Multicountry Surveys. Journal of Health, Population and Nutrition. 30 (4): 472–86

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548.

Jeong, J., Obradović, J., Rasheed, M., McCoy, D. C., Fink, G., & Yousafzai, A. K. (2019). Maternal and paternal stimulation: Mediators of parenting intervention effects on preschoolers' development. *Journal of Applied Developmental Psychology*, *60*, 105-118.

Jeong, J., McCoy, D. C., Yousafzai, A. K., Salhi, C., & Fink, G. (2016). Paternal stimulation and early child development in low-and middle-income countries. *Pediatrics*, 138(4).

Kagitcibasi, C., Sunar, D., Bekman, S., Baydar, N., & Cemalcilar, Z. (2009). Continuing effects of early enrichment in adult life: The Turkish Early Enrichment Project 22 years later. *Journal of Applied Developmental Psychology*, *30*(6), 764–779.

Kizilcec, R.F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., Turkay, S., Williams, J.J., & Tingley, D. "Scaling up behavioral science interventions in online education." *Proceedings of the National Academy of Sciences* 117, no. 26 (2020): 14900-14905.

Lebedinski, L., Carneiro, P., Arnold Urzua, T., Perng, J., Munoz Boudet, A.M., Herrera Sosa, K., 2024. "Appendix: Experimental impacts of a vritual parenting program with mothers and fathers."

Luna-Bazaldua, D. (2022). "Resultados de ECDI2030 y AIM-ECD en Serbia." [Unpublished manuscript].

Magill-Evans, J., Harrison, M. J., Rempel, G., & Slater, L. (2006). Interventions with fathers of young children: Systematic literature review. *Journal of Advanced Nursing*, *55*(2), 248–264.

Martinez, S., Naudeau, S., & Pereira, V. (2012). *The promise of preschool in Africa: A randomized impact evaluation of early childhood development in rural Mozambique*.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.

Mejia, A., Calam, R., & Sanders, M. R. (2015). A pilot randomized controlled trial of a brief parenting intervention in low-resource settings in Panama. *Prevention Science*, *16*(5), 707–717.

Mo, D., Luo, R., Liu, C., Zhang, H., Zhang, L., Medina, A., & Rozelle, S. (2014). Text messaging and its impacts on the health and education of the poor: Evidence from a field experiment in rural China. *World Development*, *64*, 766–780.

Nahar, B., Hossain, M. I., Hamadani, J. D., Ahmed, T., Huda, S. N., Grantham-McGregor, S. M., & Persson, L. A. (2012). Effects of a community-based approach of food and psychosocial stimulation on growth and development of severely malnourished children in Bangladesh: A randomised trial. *European Journal of Clinical Nutrition*, *66*(6), 701–709.

Naudeau, S., Valerio, A., Neuman, M. J., & Elder, L. K. (2011). *Investing in Young Children: An Early Childhood Development Guide for Policy Dialogue and Project Preparation*. World Bank.

Ome, A., & Menendez, A. (2021). *Using SMS and parental outreach to improve early reading skills in Zambia*. Education Economics.

Özler, B., Fernald, L. C., Kariger, P., McConnell, C., Neuman, M., & Fraga, E. (2018). Combining pre-school teacher training with parenting education: A cluster-randomized controlled trial. *Journal of Development Economics*, *133*, 448–467.

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. F. (2014). Practitioner review: Engaging fathers–recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*(11), 1187.

Peters, M. A., et al. (2022). Reimagining the new pedagogical possibilities for universities post-Covid-19: An EPAT Collective Project. *Educational Philosophy and Theory*, *54*(6), 717-760.

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin*, *128*(2), 330–366.

Richardson, B. P., Linde, J., Pillay, B., & Swanepoel, W. (2021). Do Text Messages about Health and Development in Young Children Affect Caregiver Behaviour and Child Outcomes? A Systematic Review. *Health Education Journal*, *80*(2), 225–237.

Rollè, L., Gullotta, G., Trombetta, T., Curti, L., Gerino, E., Brustia, P., & Caldarera, A. M. (2019). Father involvement and cognitive development in early and middle childhood: A systematic revie*w. Frontiers in Psychology* 10, 2405.

Rubin, D. B. (1988). An overview of multiple imputation. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 79–84.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N. and Grantham-McGregor, S., (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota. *Journal of Human Resources,* 50(2), pp.464-483.

Sayre, R., Devercelli, A., Neuman, M., & Wodon, Q. (2015). *Investing in Early Childhood Development-Review of the World Bank's Recent Experience*. World Bank.

Silander, M., Grindal, T., Hupert, N., Garcia, E., Anderson, K., Vahey, P. & Pasnik, S. (2018). What Parents Talk About When They Talk About Learning: A National Survey About Young Children and Science. New York, NY, & Menlo Park, CA: Education Development Center, Inc., & SRI International.

Singla, D. R., Kumbakumba, E., & Aboud, F. E. (2015). Effects of a parenting intervention to address maternal psychological wellbeing and child development and growth in rural Uganda: A community-based, cluster-randomised trial. *The Lancet Global Health*, *3*(8), 458–469.

Skrypnek, B. J., & Charchun, J. (2011). *An evaluation of the Nobody's Perfect parenting program*. Canadian Association of Family Resource Programs.

Sokolovic, N., Grujic, S., & Pajic, S. (2021). Evaluation of the Support, not Perfection Parenting Program in Serbia. *European Journal of Developmental Psychology*, 1–18.

Statistical Office of the Republic of Serbia. 2023. Usage ICT - households. https://www.stat.gov.rs/en-us/oblasti/upotreba-ikt/upotreba-ikt-domacinstva/

The World Bank, 2023. Serbia Policy Notes. https://documents1.worldbank.org/curated/en/0994055050423158899/pdf/IDU042c681520369 9044e709b4206703e04a8259.pdf

UNICEF, 2023. Early childhood development index 2030. https://data.unicef.org/resources/early-childhood-development-index-2030-ecdi2030/

United Nations Children's Fund, 2023. The early childhood development index 2030: A new measure of early childhood development, UNICEF. https://data.unicef.org/wp-content/uploads/2023/09/ECDI2030_Technical_Manual_Sept_2023.pdf

Vazir, S., Engle, P., Balakrishna, N., Griffiths, P. L., Johnson, S. L., Creed-Kanashiro, H., & Bentley, M. E. (2013). Cluster-randomized trial on complementary and responsive feeding education to caregivers found improved dietary intake, growth and development among rural Indian toddlers. *Maternal & Child Nutrition*, *9*(1), 99–117.

Vyrastekova, J., Huisman, J., Mosha, I., & Smits, J. (2014). Mothers more altruistic than fathers, but only when bearing responsibility alone: Evidence from parental choice experiments in Tanzania. *PLoS One*, *9*(6), e99952.

Wolf, S., Aber, J. L., Behrman, J. R., & Peele, M. (2019). Longitudinal causal impacts of preschool teacher training on Ghanaian children's school readiness: Evidence for persistence and fade-out. *Developmental Science*, *22*(5), 12878.

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), 537–566.

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

Appendix

Table 1 Meeting attendance for mothers and fathers by treatment

	Treatment 1 Standard	Treatment 2 Plus	p-value from test of equality between treatment arms
Attendance of mothers	5.6	5.1	0.01
Attendance of fathers	0.9	2.3	0.0

Note: excluding families who did not attend any sessions

Table 2 Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	**Control**	**Treatment 1**	**Treatment 2**			**p-value**	
	N=890	**N=407**	**N=238**	**N=245**	diff. all 3 groups	diff. C-T1	diff. T1-T2	diff. C-T2
Child								
Child being asked about is male	50.2%	47.9%	51.3%	53.1%	0.41	0.41	0.69	0.20
Age of child being asked about	4.3	4.3	4.4	4.3	0.22	0.09	0.28	0.58
	(1.1)	(1.1)	(1.1)	(1.0)				
Primary caregiver								
Age	32.7	32.9	32.3	32.6	0.35	0.14	0.48	0.52
	(5.6)	(5.3)	(5.6)	(5.9)				
Female	90.7%	89.7%	92.0%	91.0%	0.60	0.33	0.69	0.58
Highest education								
Primary school or less	14.8%	14.3%	13.9%	16.8%	0.60	0.89	0.37	0.38
Secondary school (VET or general)	50.2%	50.1%	52.1%	48.4%	0.71	0.63	0.41	0.66
College or higher	35.0%	35.6%	34.0%	34.8%	0.92	0.68	0.85	0.84
Employment status*								
Employed or self-employed	53.3%	51.4%	56.7%	53.1%	0.42	0.19	0.42	0.67
Student or pensioner	0.6%	0.7%	0.4%	0.4%	0.81	0.62	0.98	0.60
Unemployed, looking for job	19.3%	22.9%	16.8%	15.9%	0.05	0.07	0.79	0.03
Unemployed and not looking or homemaker	26.9%	25.1%	26.1%	30.6%	0.29	0.78	0.27	0.12
Household								
Main source of income								
Salary, self-employed, ag worker	79.9%	79.9%	82.8%	77.1%	0.30	0.36	0.12	0.41
Social assistance	12.4%	11.5%	12.6%	13.5%	0.76	0.69	0.78	0.47
Pension, investments, other sources	4.2%	4.9%	2.5%	4.5%	0.32	0.14	0.24	0.81
None	1.3%	1.2%	1.7%	1.2%	0.87	0.64	0.67	1.00
No response	2.2%	2.5%	0.4%	3.7%	0.05	0.05	0.01	0.37
Household has smartphone	90.2%	90.7%	88.7%	91.0%	0.63	0.41	0.39	0.88
Household has PC	55.6%	56.8%	56.7%	52.7%	0.55	0.99	0.37	0.31
Language								
Serbian	98.1%	98.3%	97.5%	98.4%	0.72	0.48	0.49	0.93
Hungarian	1.1%	0.5%	2.1%	1.2%	0.17	0.06	0.45	0.30
Slovak	0.8%	1.2%	0.4%	0.4%	0.39	0.30	0.98	0.29
Number of children reported in survey aged 0-7 in baseline	1.5	1.5	1.5	1.5	0.63	0.96	0.43	0.39
	(0.6)	(0.6)	(0.6)	(0.7)				

Notes: This table provides the summary statistics of the background characteristics for the overall sample and by treatment status in the columns (1) to (4). This table additionally provides p-values for the differences between the treatment arms and control group. Column (5) shows the p-values for differences by treatment status, while columns (6) through (8) show p-values for differences between control and treatment 1 (column (6)), treatment 1 and treatment 2 (column (7)) and control and treatment 2 (column (8)). Treatment 1 is standard treatment, Treatment 2 is treatment with fathers. Data are presented as mean for continuous measures with standard deviations in parenthesis (), and % for categorical measures. Source: Application questionnaire for program. *Data collected at endline.

Table 3 Impacts on child development, home environment score, number of activities with father and mother, and reports of physical punishment.

VARIABLES	(1) Child development (AIM-ECD & ECDI) score	(2) Child development (AIM-ECD 4-6 & ECDI 2-4) score	(3) Home environment score	(4) Number of activities with mother	(5) Number of activities with father	(6) Any physical punishment in the past month
Treatment	0.28***	0.28***	0.09	0.08	-0.05	-0.05
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Treatment plus fathers	-0.03	-0.01	0.12	-0.06	-0.06	-0.15*
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Observations	868	868	868	868	868	868
R-squared	0.123	0.118	0.182	0.128	0.145	0.072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of outcome	0.05	0.06	0.05	-0.01	-0.03	-0.03
SD of outcome	1.00	0.99	0.98	1.03	0.99	0.98
p-value for test of T1=T2	0.00	0.00	0.72	0.13	0.93	0.30

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.

Notes: This table shows estimates from an OLS regression. The first two columns show results of regressions of child development, an index measured by combining two instruments, the Early Childhood Development Index (ECDI 2030) and the caregiver report of Anchor Items for Measurement of Early Childhood Development (AIM-ECD). In column (1), the outcome is combined AIM-ECD and ECDI, and in column (2), we use AIM-ECD for 4-6 year olds and ECDI for 2-4 year olds. Column (3) show results from home environment score. Child's home environment is an index measured using the Home Observation Measurement of the Environment-Short Form (HOME-SF) instrument. Columns (4)-(5) show number of activities with mother (female caregiver), and the number of activities with father, respectively. Column (6) shows results from the outcome on reports of any physical punishment in the past month. Physical punishment is defined by whether the respondent shook, spanked, hit, slapped (with bare hand or object) their child in the past month. All outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: Sex/age of primary caregiver and of child being asked about; education level of primary caregiver; preferred language of caregiver; number of children at endline; implementation round of treatment; and main source of income (in baseline) and employment status of caregiver (in endline).

Table 4 Impacts on beliefs on physical punishment and on affection and attachment, generalized anxiety disorder, and beliefs about responsibility in learning.

VARIABLES	(1) Do you believe .. child needs to be physically punished?	(2) Beliefs on affection and attachment	(3) Generalized anxiety disorder (higher score more severe)	(4) A or SA do not have to worry about child learning- school responsible for learning	(5) Believes that parent most responsible for social skills	(6) Believes parent or both parent and school most responsible for academ. skills
Treatment	0.05	0.06	0.08	-0.19**	0.06	0.12
	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)
Treatment plus fathers	0.11	0.12	0.07	-0.01	0.09	0.08
	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)
Observations	855	868	868	866	867	866
R-squared	0.057	0.185	0.099	0.084	0.064	0.055
Controls	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean of outcome	0.04	0.03	0.05	-0.04	0.05	0.06
SD of outcome	1.06	0.99	1.07	0.97	0.99	0.92
p-value for test of T1=T2	0.53	0.49	0.97	0.04	0.77	0.67

Standard errors in parentheses

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

Notes: This table shows estimates from an OLS regression. The first column shows results of a regression on beliefs on necessity of physical punishment; the parent is as believing in punishment if they respond 'Yes' (versus 'No' or 'Don't know') to the question, 'Do you believe that in order to bring up, raise, or educate a child properly, the child needs to be physically punished?' Column (2) shows results from the beliefs on affection and attachment, which come from an index constructed using the Affection and Attachment factor of the Early Parenting Attitudes Questionnaire (EPAQ). Column (3) shows results from the outcome of the generalized anxiety disorderassessment (GAD-7) score; GAD-7 is a seven-item instrument. The last three columns show results of parental beliefs on who is responsible for child's learning. Column (4) shows results from the response to agreement with the statement, 'You don't have to worry about your child's learning, because he/ she will learn everything she need to know in school.' Columns (5)-(6) show estimates from two outcomes: parental beliefs on who is responsible for social skills ('Who do you think is most responsible for teaching your child about social skills, like sharing and being patient? Would you say parents are the most responsible?'), and parental beliefs on who is responsible for academic skills ('Who do you think is most responsible for teaching your child academic skills like reading, writing, and mathematics? Would you say parents are the most responsible/parent and school are equally responsible?'). All outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: Sex/age of primary caregiver and of child being asked about; education level of primary caregiver; preferred language of caregiver; number of children at endline; implementation round of treatment; and main source of income (in baseline) and employment status of caregiver (in endline).

Table 5 Romano-Wolf step-down adjusted p-values robust to multiple hypothesis testing.

	p-value for T1	p-value for T2
Numbers of activities with mother (Z-score by age)	1.00	1.00
Numbers of activities with father (Z-score by age)	1.00	1.00
Child development standarized by age	0.01	1.00
Home environment Z score, standarized by age	0.99	0.86
GAD-7 addition of score (Z-score by age)	1.00	1.00
Any physical punishment (Z-score by age)	1.00	0.68
Do you believe .. child needs to be physically punished? (Z-score by age)	1.00	0.99
Beliefs on affection (Z-score by age)	1.00	0.83
You don't have to worry about your child's learning, because (Z-score by age)	0.26	1.00
Parent most responsible for social skills (Z-score by age)	1.00	1.00
Parent or both most responsible for academ. skills (Z-score by age)	0.91	1.00

Notes: This table shows Romano-Wolf adjusted p-values robust to multiple hypothesis.

Appendix: Experimental impacts of a virtual parenting program with mothers and fathers

Lara Lebedinski Ⓡ

Pedro Carneiro Ⓡ

Tamara Arnold Urzua Ⓡ

Julie Perng Ⓡ

Ana Maria Munoz Boudet

Katia Herrera Sosa

Appendix A: Tables, and results

Dropouts

Table A1a reports the differences between parents who were initially assigned to treatment and (1) attended the program versus (2) those who dropped out from the program. We find that dropouts were more likely to be of lower education, they are more likely to have *Pension, investments, or similar* as their main source of income, they are more likely to be Hungarian and less likely to be Serbian. Overall, the dropouts have a somewhat lower socio-economic status than parents who stayed in the program.

Table A1a: Dropouts

	Total	Dropout	No dropout	p-value
	N=745	N=21	N=724	
Primary caregiver				
Highest education				
Primary school or less	21.5%	42.9%	20.9%	0.016
Secondary school (VET or general)	50.5%	52.4%	50.5%	0.86
College or higher	27.9%	4.8%	28.6%	0.016
Household				
Salary, self-employed, ag worker	75.3%	61.9%	75.7%	0.15
Social assistance	16.1%	23.8%	15.9%	0.33
Pension, investments, other sources	5.0%	14.3%	4.7%	0.046
None	1.2%	0.0%	1.2%	0.61
No response	2.4%	0.0%	2.5%	0.46
Household has smartphone	88.9%	90.5%	88.8%	0.81
Household has PC	47.9%	38.1%	48.2%	0.36
Serbian	97.3%	90.5%	97.5%	0.049
Hungarian	2.3%	9.5%	2.1%	0.024
Slovak	0.4%	0.0%	0.4%	0.77
Number of children aged 0-6	1.3 (±0.5)	1.3 (±0.5)	1.3 (±0.5)	0.89

Notes: The table compares households that dropped out with households that stayed in the program. Data are presented as mean (±SD) for continuous measures, and % for categorical measures. Source: Application questionnaire for program.

Table A1b: Differences in outcomes by educational group

	(1) Primary school or less N=132	(2) Secondary school (VET or general) N=446	(3) College or higher N=311	(4) diff. all 3 groups	(5) diff. col. (2)-(3)	(6) diff. col. (3)-(4)	(7) diff. col. (2)-(4)
					p-value		
Child development standarized by age	-0.5	0.1	0.3	<0.001	<0.001	0.001	<0.001
	(1.1)	(1.0)	(0.9)				
CD restricted AIM 4-6 & ECDI 2-4, standarized by age	-0.4	0.1	0.3	<0.001	<0.001	0.005	<0.001
	(1.1)	(1.0)	(0.9)				
Home environment Z score, standarized by age	-0.6	0.0	0.4	<0.001	<0.001	<0.001	<0.001
	(1.1)	(0.9)	(0.8)				
GAD-7 addition of score (Z-score by age)	0.4	0.1	-0.1	<0.001	0.006	0.004	<0.001
	(1.4)	(1.1)	(0.8)				
Beliefs on affection and attachment (Z-score by age)	-0.6	0.0	0.4	<0.001	<0.001	<0.001	<0.001
	(1.2)	(1.0)	(0.8)				
Do not have to worry about child's learning, because child will learn at school	0.2	-0.0	-0.1	0.014	0.030	0.27	0.004
	(1.1)	(1.0)	(0.9)				
Parent most responsible for social skills (Z-score by age)	-0.2	0.0	0.2	0.003	0.015	0.13	<0.001
	(1.0)	(1.0)	(1.0)				
Parent or both most responsible for academ. skills (Z-score by age)	0.0	0.1	0.0	0.31	0.46	0.13	0.70
	(0.9)	(0.9)	(1.0)				
Any physical punishment (Z-score by age)	0.1	-0.1	-0.0	0.085	0.028	0.59	0.091
	(1.1)	(1.0)	(1.0)				
Do you believe .. child needs to be physically punished? (Z-score by age)	-0.1	0.1	0.0	0.48	0.23	0.84	0.30
	(0.9)	(1.1)	(1.1)				
Numbers of activities with mother (Z-score by age)	-0.5	-0.0	0.2	<0.001	<0.001	0.005	<0.001
	(1.1)	(1.1)	(0.9)				
Numbers of activities with father (Z-score by age)	-0.3	-0.1	0.2	<0.001	0.10	<0.001	<0.001
	(0.9)	(1.0)	(1.0)				

Notes: This table provides the summary statistics of the outcomes by highest educational attainment of the parent in the columns (1) to (3). This table additionally provides p-values for the differences by educational attainment of the parent. Column (4) shows the p-values for differences by educational attainment overall, while columns (5) through (7) show p-values for differences between primary school or less and secondary school (column (5)), secondary school and college or higher (column (6)) and primary school or less and college or higher (column (7)). Data are presented as mean for continuous measures with standard deviations in parenthesis (), and % for categorical measures. Source: Endline survey.

Table A2: Differences in outcomes by main source of income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)

	Main source income salary, self-employed, ag worker	Main source income Social assistance (baseline)	Pension, investment, other sources	p-value			
				diff. all 3 groups	diff. col. (2)-(3)	diff. col. (3)-(4)	diff. col. (2)-(4)
	N=711	N=110	N=37				
Child development standarized by age	0.1	-0.4	-0.3	<0.001	<0.001	0.68	0.011
	(1.0)	(1.1)	(1.1)				
CD restricted AIM 4-6 & ECDI 2-4, standarized by age	0.1	-0.3	-0.2	<0.001	<0.001	0.70	0.022
	(1.0)	(1.0)	(1.1)				
Home environment Z score, standarized by age	0.2	-0.6	-0.3	<0.001	<0.001	0.11	0.002
	(0.9)	(1.1)	(1.2)				
GAD-7 addition of score (Z-score by age)	-0.0	0.5	0.1	<0.001	<0.001	0.13	0.30
	(1.0)	(1.5)	(1.0)				
Beliefs on affection and attachment (Z-score by age)	0.2	-0.5	-0.5	<0.001	<0.001	0.84	<0.001
	(0.9)	(1.2)	(1.1)				
Do not have to worry about child's learning, because child will learn at school	-0.1	0.2	0.2	0.002	0.001	0.75	0.11
	(0.9)	(1.1)	(1.1)				
Parent most responsible for social skills (Z-score by age)	0.1	-0.2	-0.2	0.001	<0.001	0.71	0.10
	(1.0)	(1.0)	(1.0)				
Parent or both most responsible for academ. skills (Z-score by age)	0.0	0.0	0.3	0.35	0.96	0.16	0.15
	(0.9)	(0.9)	(0.5)				
Any physical punishment (Z-score by age)	-0.0	0.1	-0.1	0.36	0.18	0.31	0.69
	(1.0)	(1.1)	(0.9)				
Do you believe .. child needs to be physically punished? (Z-score by age)	0.1	-0.0	-0.0	0.69	0.47	0.94	0.61
	(1.1)	(1.0)	(0.9)				
Numbers of activities with mother (Z-score by age)	0.1	-0.3	-0.1	0.007	0.002	0.29	0.54
	(1.0)	(1.1)	(1.0)				
Numbers of activities with father (Z-score by age)	0.0	-0.4	-0.3	<0.001	<0.001	0.63	0.052
	(1.0)	(1.0)	(0.9)				

Notes: This table provides the summary statistics of the outcomes by main source of income in the columns (1) to (3). This table additionally provides p-values for the differences by main income source. Column (4) shows the p-values for differences by income source overall, while columns (5) through (7) show p-values for differences between main income source salary and main income source social assistance (column (5)), main income source social assistance and pension (column (6)) and salary and pension (column (7)). Data are presented as mean for continuous measures with standard deviations in parenthesis (), and % for categorical measures. Source: Endline survey.

Table A3: Correlation matrix of outcomes

	Child dev.	Child dev.-restricted	HOME	GAD-7	Beliefs-affection	Beliefs-learning	Beliefs-social	Beliefs-academic	Phys. punishment	Beliefs-punishment	Activities mother	Activities father
Child dev.	1											
Child dev.-restricted	0.945***	1										
HOME	0.378***	0.363***	1									
GAD-7	-0.165***	-0.176***	-0.210***	1								
Beliefs-affection	0.156***	0.150***	0.256***	-0.194***	1							
Beliefs-learning	-0.0173	-0.0168	-0.103**	0.0556	-0.290***	1						
Beliefs-social	0.0556	0.0485	0.0423	-0.0193	0.133***	-0.0359	1					
Beliefs-academic	0.0302	0.0170	0.0430	0.00712	0.00531	-0.0527	-0.0321	1				
Phys. punishment	-0.224***	-0.224***	-0.197***	0.110**	-0.0626	-0.0000869	-0.0336	-0.00147	1			
Beliefs-punishment	-0.147***	-0.165***	-0.0760*	0.121***	-0.0859*	0.00632	0.000493	-0.00304	0.215***	1		
Activities mother	0.230***	0.215***	0.300***	-0.0777*	0.138***	-0.0744*	0.0555	0.0401	-0.0914**	-0.0556	1	
Activities father	0.214***	0.211***	0.330***	-0.101**	0.115***	-0.0368	0.00583	0.0238	-0.0959**	-0.0398	0.220***	1

Notes: Child dev.- Child development standardized by age, Child dev.-restricted - CD restricted AIM 4-6 & ECDI 2-4, standardized by age, HOME - Home environment Z score, standardized by age, GAD-7 - GAD-7 addition of score (Z-score by age), Beliefs-affection - Beliefs on affection and attachment (Z-score by age), Beliefs-learning - Do not have to worry about child's learning, because child will learn at school, Beliefs-social - Parent most responsible for social skills (Z-score by age), Beliefs-academic - Parent or both most responsible for academic skills (Z-score by age), Phys. punishment - Any physical punishment (Z-score by age), Beliefs-punishment - Do you believe .. child needs to be physically punished? (Z-score by age), Activities mother - Numbers of activities with mother (Z-score by age), Activities father - Numbers of activities with father (Z-score by age). *** p<0.01, ** p<0.05, * p<0.1

Table A4: Summary Statistics and Balancing Properties at Baseline

	(1) Total N=1,431	(2) Control N=686	(3) Treatment 1 N=365	(4) Treatment 2 N=380	(5) diff. all 3 groups	(6) diff. C-T1	(7) diff. T1-T2	(8) diff. C-T2
						p-value		
Not interviewed	37.8%	40.7%	34.8%	35.5%	0.10	0.062	0.83	0.10
Primary caregiver								
Highest education								
Primary school or less	14.8%	14.3%	13.9%	16.8%	0.60	0.89	0.37	0.38
Secondary school (VET or general)	50.2%	50.1%	52.1%	48.4%	0.71	0.63	0.41	0.66
College or higher	35.0%	35.6%	34.0%	34.8%	0.92	0.68	0.85	0.84
Household								
Main source of income								
Salary, self-employed, ag worker	74.1%	72.9%	76.4%	74.2%	0.46	0.21	0.48	0.64
Social assistance	17.0%	17.9%	17.3%	15.0%	0.47	0.79	0.40	0.22
Pension, investments, other sources	5.0%	5.1%	3.6%	6.3%	0.23	0.25	0.084	0.41
None	1.4%	1.6%	1.4%	1.1%	0.76	0.77	0.69	0.46
No response	2.4%	2.5%	1.4%	3.4%	0.19	0.23	0.068	0.37
Household has smartphone	88.2%	87.5%	86.8%	90.8%	0.18	0.78	0.088	0.10
Household has PC	48.8%	49.7%	47.4%	48.4%	0.77	0.48	0.78	0.69
Language								
Serbian	97.1%	96.9%	97.3%	97.4%	0.91	0.77	0.93	0.69
Hungarian	2.2%	2.0%	2.5%	2.1%	0.90	0.65	0.74	0.94
Slovak	0.7%	1.0%	0.3%	0.5%	0.34	0.18	0.59	0.40
Number of children reported in survey aged 0-6 in baseline	1.3 (0.5)	1.3 (0.6)	1.3 (0.5)	1.3 (0.5)	0.73	0.44	0.56	0.90

Notes: This table provides descriptive statistics and balancing properties at baseline for the overall sample and by treatment status in the columns (1) to (4). This table additionally provides p-values for the differences between the treatment arms and control group. Column (5) shows the p-values for differences by treatment status, while columns (6) through (8) show p-values for differences between control and treatment 1 (column (6)), treatment 1 and treatment 2 (column (7)) and control and treatment 2 (column (8)). Treatment 1 is standard treatment, Treatment 2 is treatment with fathers. Data are presented as mean for continuous measures with standard deviations in parenthesis (), and % for categorical measures. Source: Application questionnaire for program.

Table A5a: Survey Attrition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	Interviewed	Not interviewed Gr. 1	Not interviewed Gr. 2		p-value		
	N=1,431	N=890	N=382	N=159	diff. all 3 groups	diff. C-T1	diff. T1-T2	diff. C-T2
Primary caregiver								
Highest education								
Primary school or less	22.4%	14.8%	31.2%	43.9%	<0.001	<0.001	0.005	<0.001
Secondary school (VET or general)	49.2%	50.2%	49.7%	41.9%	0.16	0.89	0.10	0.058
College or higher	28.5%	35.0%	19.0%	14.2%	<0.001	<0.001	0.18	<0.001
Household								
Main source of income								
Salary, self-employed, ag worker	74.1%	79.9%	69.4%	53.5%	<0.001	<0.001	<0.001	<0.001
Social assistance	17.0%	12.4%	20.7%	34.0%	<0.001	<0.001	0.001	<0.001
Pension, investments, other sources	5.0%	4.2%	6.0%	7.5%	0.12	0.15	0.51	0.062
None	1.4%	1.3%	1.6%	1.3%	0.94	0.76	0.78	0.93
No response	2.4%	2.2%	2.4%	3.8%	0.51	0.91	0.36	0.25
Household has smartphone	88.2%	90.2%	88.7%	75.5%	<0.001	0.42	<0.001	<0.001
Household has PC	48.8%	55.6%	40.6%	30.2%	<0.001	<0.001	0.023	<0.001
Language								
Serbian	97.1%	98.1%	96.1%	94.3%	0.012	0.035	0.37	0.005
Hungarian	2.2%	1.1%	3.7%	4.4%	0.002	0.002	0.69	0.003
Slovak	0.7%	0.8%	0.3%	1.3%	0.39	0.28	0.16	0.55
Number of children reported in survey aged 0-6 in baseline	1.3	1.3	1.3	1.3	0.53	0.57	0.55	0.29
	(0.5)	(0.5)	(0.6)	(0.5)				

Notes: This table provides data on survey attrition by interviewed status in the columns (1) to (4), i.e. it compares interviewed and not interviewed households. Not interviewed households are categorized in two group. Group 1 are households who were reached but refused to participate in the survey. Group 2 are households who could not be reached by the interviewers (because they did not respond to the call, their phone number was invalid, etc.). This table additionally provides p-values for the differences by interviewed status. Column (5) shows the p-values for differences by interviewed status overall, while columns (6) through (8) show p-values for differences between interviewed and not interviewed group 1 (column (6)), not interviewed groups 1 and 2 (column (7)) and interviewed and not interviewed group 2 (column (8)). Data are presented as mean for continuous measures with standard deviations in parenthesis (), and % for categorical measures. Source: Application questionnaire for program.

Table A5b: Survey Attrition: Regression results

	(1)	(2)	(3)		(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Educational attainment				Main source of income			Assets		Language spoken at home			Number of children
	Primary school	Second. school	College or higher	Salary	Soc. ass.	Other	No source	Smart-phone	PC	Serbian	Hungarian	Slovak	0-6years
Not interviewed	0.17***	0.00	-0.17***	-0.09**	0.04	0.05**	-0.01	-0.01	-0.12**	-0.03	0.02	0.00	-0.03
	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)	(0.02)	(0.01)	(0.03)	(0.05)	(0.02)	(0.02)	(0.01)	(0.06)
Treatment	-0.00	0.02	-0.02	0.01	0.01	-0.03	0.00	-0.02	-0.00	-0.01	0.02	-0.01	-0.03
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.01)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.04)
Treatment plus fathers	0.03	-0.02	-0.01	-0.02	0.02	-0.00	0.00	0.00	-0.04	0.00	0.01	-0.01	-0.01
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.01)	(0.03)	(0.04)	(0.01)	(0.01)	(0.01)	(0.04)
Not interviewed x Treatment	0.06	-0.06	0.01	-0.09	0.12**	-0.05	0.01	-0.07*	-0.05	-0.00	0.01	-0.01	-0.00
	(0.05)	(0.07)	(0.06)	(0.06)	(0.05)	(0.03)	(0.02)	(0.04)	(0.07)	(0.02)	(0.02)	(0.01)	(0.07)
Not interviewed x Treatment plus fathers	0.03	0.02	-0.04	-0.07	0.10*	-0.02	-0.00	-0.05	-0.15**	0.02	-0.01	-0.01	-0.01
	(0.06)	(0.08)	(0.07)	(0.07)	(0.06)	(0.03)	(0.02)	(0.05)	(0.08)	(0.03)	(0.02)	(0.01)	(0.08)
Observations	1,422	1,422	1,422	1,396	1,396	1,396	1,396	1,431	1,431	1,431	1,431	1,431	1,431
R-squared	0.06	0.00	0.04	0.03	0.03	0.01	0.00	0.01	0.03	0.01	0.01	0.00	0.00

Notes: This table provides regression which tests for presence of differential selection across treatment groups. Each row shows a regression of a household characteristic on an indicator for attrition, treatment group and the interaction of the two. A statistically significant coefficient of the interaction indicates that there is differential selection across treatment groups. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. No clustering of standard errors.

Table A6: Impacts on child development *(as reported by secondary caregivers).*

VARIABLES						
	(1)	(2)	(3)	(4)	(5)	(6)
	Child development (AIM-ECD & ECDI) score (Z-score by age)			Child development (AIM-ECD 4-6 & ECDI 2-4) score (Z-score by age)		
Treatment	0.25*	0.23	0.24	0.12	0.16	0.11
	(0.14)	(0.14)	(0.15)	(0.13)	(0.13)	(0.14)
Treatment plus fathers	-0.06	-0.10	-0.06	-0.04	-0.01	0.00
	(0.14)	(0.14)	(0.15)	(0.13)	(0.13)	(0.14)
Age of child being asked about		0.06	0.04		-0.04	-0.07
		(0.06)	(0.06)		(0.05)	(0.06)
Child being asked about is male		-0.24**	-0.19		-0.05	-0.00
		(0.12)	(0.12)		(0.11)	(0.12)
Primary caregiver female		-0.30	-0.32		0.99	Jan.32
		(1.08)	(1.10)		(1.01)	(1.04)
Secondary school (VET or general)		0.45**	0.42*		0.02	0.04
		(0.21)	(0.22)		(0.20)	(0.20)
College or higher		0.61*	0.58*		0.10	0.07
		(0.33)	(0.34)		(0.31)	(0.32)
Primary caregiver unemployed, looking for job (endline)		-0.10	-0.16		0.07	0.06
		(0.17)	(0.18)		(0.16)	(0.17)
Primary caregiver not looking for job or homemaker (endline)		-0.11	-0.14		-0.15	-0.13
		(0.16)	(0.16)		(0.15)	(0.15)
Hungarian language		-0.74	-1.57*		0.02	0.12
		(0.66)	(0.88)		(0.62)	(0.83)
Slovak language		0.04	-0.29		-0.08	-0.22
		(0.94)	(0.97)		(0.88)	(0.91)
Constant	0.00	0.20	0.50	0.15*	-1.49	-1.71
	(0.08)	(1.58)	(1.66)	(0.08)	(1.48)	(1.57)
Observations	334	331	331	334	331	331
R-squared	0.015	0.084	0.155	0.004	0.088	0.145
Dep. var.: mean (std. dev)	0.05 (1.04)	0.05 (1.04)	0.05 (1.04)	0.17 (.98)	0.17 (.98)	0.17 (.98)
Controls	No	Yes	Yes	No	Yes	Yes
PI FE	No	No	Yes	No	No	Yes
Strata FE	No	Yes	Yes	No	Yes	Yes
T1=T2	0.04	0.04	0.07	0.26	0.26	0.49

Robust standard errors in parentheses

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

Controls not shown: Number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).

Table A7: *Impacts on HOME environment and GAD score (as reported by secondary caregivers).*

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Home environment score (Z-score by age)			Generalized anxiety disorder (higher score more severe) score (Z-score by age)		
Treatment	0.21	0.24*	0.27*	0.08	0.04	0.02
	(0.13)	(0.14)	(0.15)	(0.14)	(0.15)	(0.16)
Treatment plus fathers	0.15	0.11	0.15	-0.06	-0.10	-0.12
	(0.13)	(0.14)	(0.15)	(0.15)	(0.15)	(0.16)
Age of child being asked about		-0.01	-0.02		0.04	0.08
		(0.05)	(0.06)		(0.06)	(0.06)
Child being asked about is male		0.02	0.03		-0.09	-0.14
		(0.11)	(0.12)		(0.12)	(0.13)
Primary caregiver female		-0.59	-0.56		-4.12***	-4.20***
		(1.03)	(1.08)		(1.12)	(1.16)
Secondary school (VET or general)		-0.19	-0.26		-0.40*	-0.38*
		(0.20)	(0.21)		(0.22)	(0.23)
College or higher		0.25	0.24		-0.35	-0.34
		(0.31)	(0.33)		(0.34)	(0.36)
Primary caregiver unemployed, looking for job (endline)		0.11	0.10		0.03	0.06
		(0.16)	(0.17)		(0.18)	(0.18)
Primary caregiver not looking for job or homemaker (endline)		0.07	0.05		0.03	0.07
		(0.15)	(0.16)		(0.16)	(0.17)
Hungarian language		-0.08	0.45		-0.62	0.16
		(0.63)	(0.87)		(0.69)	(0.93)
Slovak language		0.93	0.68		0.75	1.00
		(0.89)	(0.95)		(0.97)	(1.02)
Constant	0.00	2.28	2.30	0.00	4.12**	4.20**
	(0.08)	(1.51)	(1.63)	(0.09)	(1.65)	(1.75)
Observations	331	328	328	334	331	331
R-squared	0.008	0.108	0.130	0.002	0.116	0.163
Dep. var.: mean (std. dev)	0.1 (1.01)	0.1 (1.01)	0.1 (1.01)	0.01 (1.1)	0.01 (1.1)	0.01 (1.1)
Controls	No	Yes	Yes	No	Yes	Yes
PI FE	No	No	Yes	No	No	Yes
Strata FE	No	Yes	Yes	No	Yes	Yes
T1=T2	0.69	0.38	0.45	0.37	0.40	0.41

Robust standard errors in parentheses

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

Controls not shown: Number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).

Table A8: Impacts on beliefs on affection and attachment (as reported by secondary caregivers).

VARIABLES	(1)	(2)	(3)
		Beliefs on affection and attachment	
Treatment	-0.03	0.01	0.07
	(0.14)	(0.14)	(0.15)
Treatment plus fathers	0.05	0.07	0.09
	(0.14)	(0.14)	(0.15)
Age of child being asked about		-0.09	-0.09
		(0.06)	(0.06)
Child being asked about is male		0.12	0.20*
		(0.12)	(0.12)
Primary caregiver female		0.25	0.27
		(1.09)	(1.09)
Secondary school (VET or general)		0.07	-0.10
		(0.21)	(0.21)
College or higher		-0.11	-0.15
		(0.33)	(0.34)
Primary caregiver unemployed, looking for job (endline)		0.06	0.05
		(0.17)	(0.17)
Primary caregiver not looking for job or homemaker (endline)		-0.15	-0.12
		(0.16)	(0.16)
Hungarian language		0.49	0.46
		(0.67)	(0.87)
Slovak language		1.22	0.65
		(0.94)	(0.96)
Constant	-0.00	1.30	0.28
	(0.09)	(1.60)	(1.64)
Observations	334	331	331
R-squared	0.001	0.105	0.206
Dep. var.: mean (std. dev)	0.01 (1.07)	0.01 (1.07)	0.01 (1.07)
Controls	No	Yes	Yes
PI FE	No	No	Yes
Strata FE	No	Yes	Yes
T1=T2	0.61	0.75	0.91

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Controls not shown: Number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).

Table A9: Impacts on physical punishment behaviors and attitudes (as reported by secondary caregivers).

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Any physical punishment in the past month			Do you believe a child needs to be physically punished?		
Treatment	0.01	0.04	0.04	-0.07	-0.02	0.03
	(0.13)	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)
Treatment plus fathers	-0.09	-0.12	-0.11	0.10	0.07	0.04
	(0.13)	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)
Age of child being asked about		0.02	0.00		-0.01	0.02
		(0.05)	(0.06)		(0.06)	(0.06)
Child being asked about is male		0.16	0.17		0.23*	0.23*
		(0.11)	(0.12)		(0.12)	(0.13)
Primary caregiver female		-2.42**	-2.20**		0.84	0.42
		(1.03)	(1.07)		(1.09)	(1.11)
Secondary school (VET or general)		-0.06	-0.06		-0.14	-0.14
		(0.20)	(0.21)		(0.21)	(0.22)
College or higher		-0.15	-0.11		0.20	0.28
		(0.31)	(0.33)		(0.33)	(0.34)
Primary caregiver unemployed, looking for job (endline)		-0.07	-0.10		-0.09	-0.05
		(0.16)	(0.17)		(0.17)	(0.18)
Primary caregiver not looking for job or homemaker (endline)		-0.08	-0.09		-0.05	0.01
		(0.15)	(0.16)		(0.16)	(0.16)
Hungarian language		0.17	0.79		1.00	1.65*
		(0.63)	(0.86)		(0.74)	(0.95)
Slovak language		0.15	0.08		0.15	0.19
		(0.89)	(0.94)		(0.96)	(0.99)
Constant	-0.00	Jan.74	Jan.76	0.00	-2.21	-2.00
	(0.08)	(1.51)	(1.62)	(0.09)	(1.72)	(1.81)
Observations	334	331	331	329	326	326
R-squared	0.002	0.065	0.106	0.003	0.069	0.144
Dep. var.: mean (std. dev)	-0.02 (.98)	-0.02 (.98)	-0.02 (.98)	0.01 (1.04)	0.01 (1.04)	0.01 (1.04)
Controls	No	Yes	Yes	No	Yes	Yes
PI FE	No	No	Yes	No	No	Yes
Strata FE	No	Yes	Yes	No	Yes	Yes
T1=T2	0.47	0.27	0.34	0.29	0.59	0.94

Robust standard errors in parentheses

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

Controls not shown: Number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).

Table A10: Impacts on number of activities with male caregivers (as reported by secondary caregivers).

VARIABLES	(1)	(2)	(3)
	Number of activities with father (Z-score by age)		
Treatment	0.14	0.08	0.08
	(0.14)	(0.15)	(0.16)
Treatment plus fathers	0.05	0.01	0.03
	(0.14)	(0.15)	(0.16)
Age of child being asked about		0.11*	0.12*
		(0.06)	(0.06)
Child being asked about is male		-0.30**	-0.35***
		(0.12)	(0.13)
Primary caregiver female		0.16	-0.13
		(1.13)	(1.17)
Secondary school (VET or general)		-0.09	-0.11
		(0.22)	(0.23)
College or higher		-0.01	0.08
		(0.34)	(0.36)
Primary caregiver unemployed, looking for job (endline)		0.02	-0.04
		(0.18)	(0.19)
Primary caregiver not looking for job or homemaker (endline)		-0.08	-0.18
		(0.16)	(0.17)
Hungarian language		0.30	0.76
		(0.69)	(0.93)
Slovak language		0.54	-0.02
		(0.98)	(1.03)
Constant	-0.00	0.77	0.86
	(0.09)	(1.65)	(1.76)
Observations	331	328	328
R-squared	0.003	0.060	0.107
Dep. var.: mean (std. dev)	0.05 (1.08)	0.05 (1.08)	0.05 (1.08)
Controls	No	Yes	Yes
PI FE	No	No	Yes
Strata FE	No	Yes	Yes
T1=T2	0.57	0.70	0.77

Robust standard errors in parentheses

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

Controls not shown: Number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).

Table A11: Results for household environment-related outcomes after multiple imputation

VARIABLES	(1) Numbers of activities with mother (Z-score by age)	(2) Numbers of activities with father (Z-score by age)	(3) Any physical punishment (Z-score by age)	(4) Do you believe .. child needs to be physically punished? (Z-score by age)	(5) Home environment Z score standardized by age
Treatment	0.06	-0.02	-0.11	0.06	0.06
	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)
Treatment plus fathers	-0.03	-0.03	-0.13*	0.11	0.07
	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)
Constant	0.74	1.01	0.26	0.22	-0.17
	(0.89)	(0.80)	(0.78)	(0.87)	(0.86)
Observations	1,379	1,379	1,379	1,378	1,379
Controls	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
T1=T2	0.35	0.96	0.87	0.64	0.85

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A12: Results for beliefs-related outcomes after multiple imputation

VARIABLES	(1) GAD-7 addition of score (Z-score by age)	(2) Beliefs on affection and attachment (Z-score by age)	(3) Do not have to worry about child's learning because child will learn at school	(4) Parent most responsible for social skills (Z-score by age)	(5) Parent or both most responsible for academ. skills (Z-score by age)
Treatment	0.13	0.02	-0.08	0.03	0.05
	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)
Treatment plus fathers	0.09	0.04	0.02	-0.01	0.03
	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)
Constant	-0.85	-1.59**	1.73**	-1.40*	0.29
	(0.80)	(0.73)	(0.79)	(0.83)	(0.72)
Observations	1,379	1,379	1,379	1,379	1,379
Controls	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
T1=T2	0.71	0.81	0.27	0.68	0.78

Standard errors in parentheses

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

Table A13: Results for child development-related outcomes after multiple imputation

VARIABLES	(1) Child development standardized by age	(2) CD restricted AIM 4-6 & ECDI 2-4 standardized by age	(3) AIM-ECD standardized by age	(4) ECDI standardized by age
Treatment	0.17**	0.16**	0.15*	0.18**
	(0.08)	(0.08)	(0.08)	(0.08)
Treatment plus fathers	-0.05	-0.02	-0.05	-0.06
	(0.08)	(0.08)	(0.08)	(0.08)
Constant	-0.63	-0.24	-0.46	-0.68
	(0.81)	(0.80)	(0.83)	(0.79)
Observations	1,379	1,379	1,379	1,379
Controls	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
T1=T2	0.01	0.05	0.03	0.01

Standard errors in parentheses

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

Appendix B: Heterogeneous impacts:

Table B1 Heterogeneous impacts on child development.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Child development standardized by age			AIM-ECD standardized by age			ECDI standardized by age		
Treatment	0.27**	0.36***	0.24**	0.20*	0.38***	0.22*	0.27**	0.26**	0.22*
	(0.11)	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)
Treatment plus fathers	-0.19	-0.09	0.05	-0.23**	-0.12	0.02	-0.12	-0.02	0.09
	(0.12)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)
Treatment x male child	0.03			0.17			-0.10		
	(0.16)			(0.16)			(0.17)		
Treatment plus fathers x male child	0.31*			0.37**			0.22		
	(0.16)			(0.16)			(0.16)		
Child being asked about is male	-0.35***			-0.39***			-0.29***		
	(0.10)			(0.10)			(0.10)		
Treatment x college education		-0.21			-0.28*			-0.12	
		(0.17)			(0.17)			(0.17)	
Treatment plus fathers x college education		0.19			0.21			0.05	
		(0.17)			(0.17)			(0.17)	
College or higher		0.54***			0.56***			0.48**	
		(0.20)			(0.20)			(0.20)	
Treatment x Older child (age 5 to 6)			0.15			0.19			0.05
			(0.15)			(0.15)			(0.15)
Treatment plus fathers x Older child (age 5 to 6)			-0.10			-0.08			-0.15
			(0.15)			(0.15)			(0.15)
Older child (age 5 to 6)			0.04			0.04			0.05
			(0.10)			(0.10)			(0.10)
Constant	-0.29	-0.33	-0.41	-0.35	-0.39	-0.44	-0.11	-0.16	-0.44
	(0.67)	(0.67)	(0.66)	(0.67)	(0.67)	(0.67)	(0.68)	(0.68)	(0.67)
Observations	868	868	868	868	868	868	868	868	868
R-squared	0.127	0.127	0.117	0.125	0.127	0.122	0.096	0.093	0.094
Dep. var.: mean (std. dev)	0.05 (1)	0.05 (1)	0.05 (1)	0.05 (1)	0.05 (1)	0.05 (1)	0.04 (1)	0.04 (1)	0.04 (1)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
T2=T3	0.00	0.00	0.12	0.00	0.00	0.11	0.00	0.01	0.31

Notes: This table shows heterogenous impacts on different child development outcomes by child gender, by educational attainment of the reporting caregiver and by age group of the child. The outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: age of child being asked about, gender of the child, gender of primary caregiver, educational level of primary caregiver, primary caregiver student or pensioner (endline), primary caregiver unemployed, looking for job (endline), primary caregiver not looking for job or homemaker (endline), Hungarian language, Slovak language, number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No clustering of standard errors.

Table B2 Heterogeneous impacts on beliefs on affection, learning, social skills, and academic skills.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Beliefs on affection and attachment (Z-score by age)			A or SA don't have to worry about your child's learning- school responsible for child's learning			A or SA parent most responsible for social skills (Z-score by age)			A or SA parent or both most responsible for academic skills (Z-score by age)		
Treatment	0.03	0.08	0.05	-0.14	-0.11	-0.14	0.03	0.03	0.09	0.21*	0.03	0.12
	(0.11)	(0.09)	(0.11)	(0.11)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)	(0.11)	(0.09)	(0.11)
Treatment plus fathers	0.13	0.14	0.24**	-0.05	0.11	-0.08	0.09	0.06	0.14	0.17	0.08	0.02
	(0.11)	(0.10)	(0.10)	(0.11)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.10)
Treatment x male child	0.07			-0.10			0.06			-0.19		
	(0.15)			(0.16)			(0.17)			(0.15)		
Treatment plus fathers x male child	-0.02			0.08			-0.00			-0.18		
	(0.15)			(0.16)			(0.17)			(0.15)		
Child being asked about is male	-0.04			-0.07			-0.02			0.21**		
	(0.09)			(0.10)			(0.10)			(0.10)		
Treatment x college education		-0.04			-0.21			0.09			0.25	
		(0.16)			(0.17)			(0.17)			(0.16)	
Treatment plus fathers x college education		-0.05			-0.33**			0.08			-0.01	
		(0.16)			(0.17)			(0.17)			(0.16)	
College or higher		0.46**			0.17			0.11			-0.20	
		(0.19)			(0.19)			(0.20)			(0.19)	
Treatment x Older child (age 5 to 6)			0.03			-0.11			-0.07			-0.01
			(0.15)			(0.16)			(0.17)			(0.15)
Treatment plus fathers x Older child (age 5 to 6)			-0.25			0.16			-0.12			0.14
			(0.16)			(0.16)			(0.17)			(0.16)
Older child (age 5 to 6)			0.08			-0.02			0.03			-0.05
			(0.10)			(0.10)			(0.10)			(0.10)
Constant	-1.59**	-1.60**	-1.40**	1.72***	1.62**	1.57**	-1.22*	-1.20*	-1.26*	0.02	0.06	0.07
	(0.64)	(0.64)	(0.63)	(0.66)	(0.66)	(0.65)	(0.69)	(0.69)	(0.68)	(0.64)	(0.64)	(0.63)
Observations	868	868	868	866	866	868	867	867	867	866	866	866
R-squared	0.186	0.185	0.185	0.085	0.089	0.085	0.064	0.065	0.065	0.057	0.058	0.056
	0.03	0.03	0.03	-0.03	-0.03	-0.03	0.05	0.05	0.05	0.06	0.06	0.06
Dep. var.: mean (std. dev)	(0.99)	(0.99)	(0.99)	(0.97)	(0.97)	(0.97)	(0.99)	(0.99)	(0.99)	(0.92)	(0.92)	(0.92)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
T2=T3	0.41	0.56	0.12	0.48	0.04	0.65	0.66	0.79	0.69	0.75	0.61	0.38

Notes: This table shows heterogenous impacts on parental beliefs by child gender, by educational attainment of the reporting caregiver and by age group of the child. The outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: age of child being asked about, gender of the child, gender of primary caregiver, educational level of primary caregiver, primary caregiver student or pensioner (endline), primary caregiver unemployed, looking for job (endline), primary caregiver not looking for job or homemaker (endline), Hungarian language, Slovak language, number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline).. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No clustering of standard errors.

Table B3 Heterogeneous impacts on the HOME environment and GAD score.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Home environment score (Z-score by age)			GAD-7 addition of score (Z-score by age)		
Treatment	-0.02	0.11	0.07	0.17	0.10	0.01
	(0.11)	(0.09)	(0.10)	(0.12)	(0.11)	(0.12)
Treatment plus fathers	0.05	0.13	0.27***	0.22*	0.16	0.02
	(0.11)	(0.09)	(0.10)	(0.13)	(0.11)	(0.12)
Treatment x male child	0.22			-0.19		
	(0.15)			(0.18)		
Treatment plus fathers x male child	0.13			-0.29*		
	(0.15)			(0.18)		
Child being asked about is male	-0.18*			0.18		
	(0.09)			(0.11)		
Treatment x college education		-0.07			-0.06	
		(0.16)			(0.18)	
Treatment plus fathers x college education		-0.03			-0.24	
		(0.16)			(0.18)	
College or higher		0.79***			-0.02	
		(0.18)			(0.21)	
Treatment x Older child (age 5 to 6)			0.03			0.16
			(0.15)			(0.18)
Treatment plus fathers x Older child (age 5 to 6)			-0.33**			0.12
			(0.15)			(0.18)
Older child (age 5 to 6)			0.13			-0.09
			(0.09)			(0.11)
Constant	-1.17*	-1.21*	-1.18*	-1.41*	-1.38*	-1.25*
	(0.63)	(0.63)	(0.62)	(0.73)	(0.73)	(0.72)
Observations	868	868	868	868	868	868
R-squared	0.184	0.182	0.188	0.103	0.101	0.100
Dep. var.: mean (std. dev)	0.05 (0.98)	0.05 (0.98)	0.05 (0.98)	0.05 (1.07)	0.05 (1.07)	0.05 (1.07)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes
T2=T3	0.54	0.90	0.09	0.72	0.61	0.91

Notes: This table shows heterogenous impacts on home environment and anxiety by child gender, by educational attainment of the reporting caregiver and by age group of the child. The outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: age of child being asked about, gender of the child, gender of primary caregiver, educational level of primary caregiver, primary caregiver student or pensioner (endline), primary caregiver unemployed, looking for job (endline), primary caregiver not looking for job or homemaker (endline), Hungarian language, Slovak language, number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No clustering of standard errors.

Table B4 Heterogeneous impacts on activities with male/female caregivers and on physical punishment.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Numbers of activities with mother (Z-score by age)			Numbers of activities with father (Z-score by age)			Any physical punishment (Z-score by age)			Do you believe .. child needs to be physically punished? (Z-score by age)		
Treatment	0.03	-0.02	0.03	-0.18	-0.03	-0.04	-0.10	-0.03	-0.14	0.05	-0.03	0.10
	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.13)	(0.11)	(0.12)
Treatment plus fathers	-0.01	-0.20*	-0.00	-0.05	-0.13	0.02	-0.18	-0.06	-0.19*	0.19	0.16	0.14
	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.11)	(0.12)	(0.10)	(0.11)	(0.13)	(0.11)	(0.12)
Treatment x male child	0.09			0.25			0.10			-0.00		
	(0.17)			(0.16)			(0.16)			(0.18)		
Treatment plus fathers x male child	-0.10			-0.01			0.06			-0.15		
	(0.17)			(0.16)			(0.16)			(0.18)		
Child being asked about is male	-0.19*			-0.17*			0.29***			0.21*		
	(0.10)			(0.10)			(0.10)			(0.11)		
Treatment x college education		0.28			-0.06			-0.08			0.22	
		(0.17)			(0.16)			(0.17)			(0.19)	
Treatment plus fathers x college education		0.39**			0.22			-0.25			-0.13	
		(0.17)			(0.16)			(0.17)			(0.19)	
College or higher		0.24			0.16			-0.24			-0.05	
		(0.20)			(0.19)			(0.20)			(0.22)	
Treatment x Older child (age 5 to 6)			0.11			-0.02			0.19			-0.13
			(0.17)			(0.16)			(0.16)			(0.18)
Treatment plus fathers x Older child (age 5 to 6)			-0.13			-0.17			0.09			-0.07
			(0.17)			(0.16)			(0.17)			(0.18)
Older child (age 5 to 6)			0.06			0.11			-0.07			-0.01
			(0.10)			(0.10)			(0.10)			(0.11)
Constant	0.09	0.21	0.20	1.01	1.08*	1.13*	-0.07	-0.15		0.33	0.34	0.19
	(0.69)	(0.68)	(0.68)	(0.65)	(0.65)	(0.64)	(0.68)	(0.68)		(0.74)	(0.74)	(0.73)
Observations	868	868	868	868	868	868	868	868		855	855	855
R-squared	0.129	0.134	0.129	0.149	0.148	0.146	0.073	0.075		0.058	0.061	0.056
Dep. var.: mean (std. dev)	-0.01	-0.01	-0.01	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	0.03		0.03
	(1.03)	(1.03)	(1.03)	(0.99)	(0.99)	(0.99	(0.98)	(0.98)	(0.98)	(1.06)		(1.06)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PI FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
T2=T3	0.75	0.12	0.80	0.32	0.33		0.57	0.76	0.72	0.33	0.13	0.79

Notes: This table shows heterogenous impacts on number of activities with mother and father and physical punishment by child gender, by educational attainment of the reporting caregiver and by age group of the child. The outcomes have been standardized within age, using the means and standard deviations in the control group. Controls not shown: age of child being asked about, gender of the child, gender of primary caregiver, educational level of primary caregiver, primary caregiver student or pensioner (endline), primary caregiver unemployed, looking for job (endline), primary caregiver not looking for job or homemaker (endline), Hungarian language, Slovak language, number of children at endline, age of primary caregiver, implementation round of treatment, and main source of income (in baseline). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No clustering of standard errors.

Appendix C: Intervention locations:

Regions	District	Municipality
Vojvodina, Belgrade region, Region of South-East Serbia and Region of Sumadia and West Serbia	Belgrade, Bor District, Braničevo district, Jablanica District, Kolubara district, Nišava district, Pčinja District, Podunavlje district, South Bačka district, South Banat district, Šumadijski district, West Bačka district, Zlatibor district	(Bač, Bečej, Kragujevac, Niš, Odžaci, Pančevo, Požarevac, Sjenica, Smederevska Palanka, Svrljig, Valjevo, Vranje and Žabalj, Apatin, Golubac, Lazarevac, Majdanpek, Palilula, Smederevo, Velika Plana and Vlasotince